

Europe's arrested development

Encouraging statistics on industrial R&D don't tell the whole story.

Major companies in Europe are at last spending more on research and development (R&D), according to figures released earlier this month. The news is a welcome, albeit belated, indication that they are responding to the challenges posed by their global competitors.

The largest 1,000 companies in the European Union (EU) increased their total R&D spending by 5.3% last year, the preliminary data released by the European Commission reveal. This spending will not, in itself, result in the level of innovation that European industry requires to remain competitive. It does, however, indicate that one element of the overall picture — large-scale industrial research — is moving in the right direction.

But the Industrial R&D Investment Scoreboard doesn't tell the whole story. For a start, although it tracks the total R&D activity of European companies, a large and growing proportion of this takes place outside the EU. The scoreboard also found that the world's top 1,000 companies grew their research even more quickly — by 7.7% — during 2005. And the officials that compiled the data caution that, despite the growth, European industry still faces structural issues that do not auger well for the future of innovation in the EU.

Political leaders now accept unequivocally that research, development and innovation are critical to Europe's competitiveness, and therefore to the future of the EU. Four years ago, at a summit in Barcelona, they agreed that the EU should attempt to increase its total investment in R&D to 3% of its economy.

R&D as a percentage of economic output is not an ideal metric, and many economists question whether 3% is an attainable target. Nonetheless, it provides a useful yardstick against which progress can be measured, as it includes total R&D carried out, not just government expenditure. That's important, as industrial R&D in Europe is much weaker than in the United States or Japan, and will probably have to strengthen if the continent is to sustain a decent level of economic growth. Not to put too fine a point on it, innovation is the only thing that can prevent European industry being inexorably squeezed between the talent and power of the United States and the rapid industrialization in Asia.

The latest data show, for the first time in the scoreboard's brief

three-year history, that research at large European companies has been expanding substantially. Other data, obtained by a European Commission survey of companies' investment plans, indicate that this improvement is likely to be sustained, at least in the medium term. These findings are encouraging.

Both the scorecard and the survey deal with large companies, which between them account for most industrial research. But it is well established that innovation in small companies — or rather, the lack of it — is a critical issue in Europe. Here, too, there are some encouraging signs. There is anecdotal evidence that the old model of European small- and medium-sized firms — as family-owned and intensely conservative — is gradually shifting. According to European Commission officials, a study of employment data suggests that more scientists and engineers are now working for small companies.

But large firms remain a major driver of industrial innovation — and here the outlook is mixed. Sectors of established European strength, such as cars and chemicals, are huge spenders on product development, but they are not particularly research intensive. The sectors that are most research intensive — and that are the main focal points for global innovation — are pharmaceuticals and the broad sphere of activity known these days as information and communications technology (ICT).

The news there is not so good. Many factors contribute to Europe's poor track record in ICT, including the continued fragmentation of the continent's 'common market' and the lack of an innovation engine in any way comparable to California's Silicon Valley. Pharmaceuticals research and the biotechnology industry are both increasingly concentrated in the United States.

The R&D scoreboard, which will be published in full later in the year, makes a valuable contribution to this discussion by highlighting some of Europe's strengths and weaknesses. Political discourse in Europe currently tends to focus on the expansion of the European Union, and on social issues, such as immigration. Scientists, engineers and business leaders should miss no opportunity to remind their governments of the 3% target, and to highlight the importance of industrial innovation. Without it, Europe faces a future on the economic sidelines. ■

Rich in plutonium

The US nuclear-weapons complex is too large — and is likely to remain so.

In the wake of North Korea's nuclear-weapons test, little attention has been paid to a revised plan for US nuclear weapons. But details of the scheme emerged on 19 October when officials announced the start of its environmental review. The plan, which focuses heavily

on the handling of the plutonium used in nuclear weapons, deserves close scrutiny: it could have wide implications not just for the United States' own nuclear-weapons stockpile, but for the fate of the tattered non-proliferation regime.

The United States currently retains about 50 tonnes of plutonium for military use — enough to fuel some 9,000 warheads. Where it isn't deployed, the plutonium is scattered across the nation's sprawling nuclear-weapons complex.

Last week, officials at the National Nuclear Security Administration (NNSA), the branch of the US Department of Energy that runs the



nuclear-weapons programme, announced that it would like to reduce costs by moving nearly all its plutonium into a single facility. The consolidated plutonium centre, as it would be called, would be built at a pre-existing site and would place plutonium-related research, surveillance and manufacturing under a single roof.

That would be a positive move. Plutonium is too widely dispersed under the current arrangement, with significant quantities being held near populated areas, such as Livermore in California. Confining plutonium to a single, remote site therefore makes sense — as do elements of the plan that would consolidate existing high-explosive and hydrodynamic testing facilities.

Other aspects of the plan are more troubling, however. Specifically, the plutonium centre would be equipped to produce around 125 new plutonium 'pits' per year — these are the cores of modern nuclear weapons. NNSA officials say that the new production capability is necessary to help build a "reliable replacement warhead" — a new kind of warhead that is supposed to require less maintenance than existing designs (see *Nature* 442, 18–21; 2006). They argue that these warheads would replace, rather than augment, the existing stockpile.

But the construction of such a facility is liable to prompt other nations to revisit their own production plans. Although it won't

change the course of countries such as Iran, it reinforces the increasingly prevalent view that nuclear weapons are vital to any nation's security. That perception may encourage further development in non-nuclear states, such as Japan and Brazil. It also sends a clear message to the United States' old nemesis Russia, which continues to maintain an unnecessarily large nuclear stockpile of its own.

The United States has not yet made a convincing case that the new warheads are needed — tests on existing plutonium pits suggest that they will last for several decades, and the replacement warhead is in any case still in the design phase. What's more, the United States already has a working pit facility at Los Alamos National Laboratory in New Mexico, which some specialists think could be adapted to produce some 50 pits a year.

Increasing pit production and building new kinds of weapons run contrary to the spirit of the 1970 Nuclear Non-Proliferation Treaty, which calls for existing nuclear-weapons states to take steps towards disarmament. Instead, the United States and other nuclear states should be acting to shrink both their weapons stockpiles and their production complexes. The best thing they could do for their collective future security would be to demonstrate to the world that nuclear weapons are less central to their own defence strategies than they were during the cold war. ■

Plan bee

Another day, another genome.

Today, *Nature* publishes a genome sequence and analysis of the western honeybee, *Apis mellifera* (see pages 893, 919 and 931). But hasn't this genome lark become a bit ho-hum? From dogs to trees to microbes and, of course, people, the US National Human Genome Research Institute lists more than 50 genome projects either complete or under way. The publication of a genome sequence used to be so exciting. Is it now destined to be dull?

Not if supporters of the honeybee project have anything to say about it. Two other insects have already been sequenced: the malaria-carrying mosquito *Anopheles gambiae*, and one of science's great model organisms, the fruitfly *Drosophila melanogaster*. Like these, the bee is much easier to manipulate and study than, say, the monkey. But unlike the mosquito and the fruitfly, the bee's social behaviour is of special interest.

Many schoolchildren learn about the waggle dance, the remarkable figure-of-eight movement that bees use to pass on the location of food. And children aren't the only ones entranced by the hierarchy of bees. Every bee has its place, from the nurses who tend the young to the foragers who collect food. There is something amazing about this insect's ability to self-organize into a productive, harmonious unit.

Indeed, the bee may fascinate us precisely because of the apparent parallels between bee society and our own — from the humble worker to the regal queen. Bee jobs are highly specialized: there are forager bees, for instance, who only scout for food, and those who follow the scouts' lead once food is found. There are bees who prefer to gather nectar and others who choose pollen. There are even

designated undertaker bees, who remove their dead comrades from the colony. This tidy division of labour mirrors what we see in our own species every day. And in perhaps the strongest parallel of all, the bee hierarchy sometimes breaks down entirely — a situation aptly termed 'anarchy'.

It is tempting to wonder whether the mechanisms controlling such complex bee behaviour are related to those at work in human society. The bee genome may allow us to get at the roots of social behaviour that is similar to our own. Researchers cannot engineer human society, by separating families, for example, or setting up entire communities in which every inhabitant is the same age. But they can do this in bees — and then measure the effects of these interventions on bee genes and physiology. And that can help elucidate the environmental, genetic and social roots of bee behaviour. Herein lies the utility of the bee: it lies somewhere between the human and the fruitfly. Like us, it is both complex and familiar. But like a fruitfly, it is easy to study in controlled conditions.

Of course, the comparison between bees and humans can only go so far. Bee individuals in a colony are more closely related than people in most cities or towns. And although there are some genetic and mechanical similarities between people and bees, we are vastly different creatures. We don't know what drives most bee or human behaviours, so it's anyone's guess whether they're at all related. But for the community of bee researchers who study this insect purely because it is an intriguing animal in its own right, that doesn't really matter. Regardless of whether the bee tells us anything about ourselves, it fully deserves its moment in the sun. ■

"There is something amazing about this insect's ability to self-organize into a productive, harmonious unit."

RESEARCH HIGHLIGHTS

Hear, hear

Cell **127**, 277–289 (2006)

Researchers have uncovered a novel mechanism underlying inherited deafness.

Christine Petit of the Pasteur Institute in Paris, France, and her colleagues studied the mouse equivalent of a protein known to be defective in some people who are profoundly deaf. They found that the protein, otoferlin, is sited at a key location within the inner hair cells (pictured) of the cochlea.

These cells transform sound into signals that trigger auditory nerves to fire. Sacs of neurotransmitters are anchored to the inner side of membranes of these hair cells. They fuse with the membrane to release their contents, activating neighbouring nerve endings. Otoferlin is essential for fusion.



S. GSCHMEISSNER/SPL

PLANETARY SCIENCE

Frosted Earths

Astrophys. J. **650**, L139–L142 (2006)

Recent observations have shown that some small stars called M dwarfs host icy planets that are roughly ten times more massive than the Earth. How are these 'super-Earths' made?

Planet formation around dwarf stars is different to that around Sun-like stars. This is because the dwarfs fade and shrink during the process, pulling in the 'snow line', which separates regions of icy-planet formation from those of rocky-planet formation.

Grant Kennedy of the Australian National University in Weston Creek and his team have concocted a theoretical model of this process, showing that it favours the rapid appearance of middleweight icy planets. As the contracting snow line moves like a cold front over small rocky protoplanets, they become thickly ice-coated before coagulating through collisions to make super-Earths.

DRUG DISCOVERY

Target practice

Proc. Natl Acad. Sci. USA **103**, 15422–15427 (2006)

An analysis of how one small molecule interrupts a protein–protein interaction may help researchers to design new drugs.

Protein–protein interactions are promising drug targets, but researchers have struggled to find footholds for small molecules in flat protein–protein interfaces. Jim Wells, then at Sunesis Pharmaceuticals in San Francisco, California, and his colleagues studied a small molecule that blocks the interaction of two proteins — IL-2R α and

IL-2, involved in conveying immune signals.

This small molecule binds within a crevice of IL-2 (pictured below). They found that it targets many of the same contact points as does IL-2R α , despite having a different structure. This is possible because IL-2 is very flexible. The finding shows that small molecules do not need to structurally mimic the proteins they displace.

STEM CELLS

Grown naturally

Nature Biotechnol. doi:10.1038/nbt1259 (2006)

Researchers have edged a step closer to making cells that might cure diabetes.

Diabetes occurs when 'beta' cells in the human pancreas fail to make enough of the hormone insulin. Making functional beta cells from human embryonic stem cells might cure this deficit, but it has proved difficult.

A team led by Emmanuel Baetge at the biotechnology company Novocell in San Diego, California, approached

the problem by trying to coax human embryonic stem cells through the stages of normal fetal pancreatic development. The stem cells did develop into cells that produce high levels of insulin, but not in response to the body's normal chemical triggers.

CELL BIOLOGY

A chaperone for arsenic

Proc. Natl Acad. Sci. USA **103**, 15617–15622 (2006)

Arsenic is flushed through biological systems with the help of a protein that clings to the toxic metal and guides it to a cellular-scale pump, a new study finds.

Researchers are driven to understand arsenic toxicity because the metal contaminates water supplies in areas such as Bangladesh and West Bengal. In this study, Barry Rosen of Wayne State University in Detroit, Michigan, Adrian Walmsley of Durham University, UK, and their team identify a protein, ArsD, in bacterial cells that picks up arsenite ions from the cell's cytoplasm. ArsD then liaises with an enzyme to activate the cell's efflux pump. The protein is therefore acting as a metallochaperone — the first to be described for arsenic.

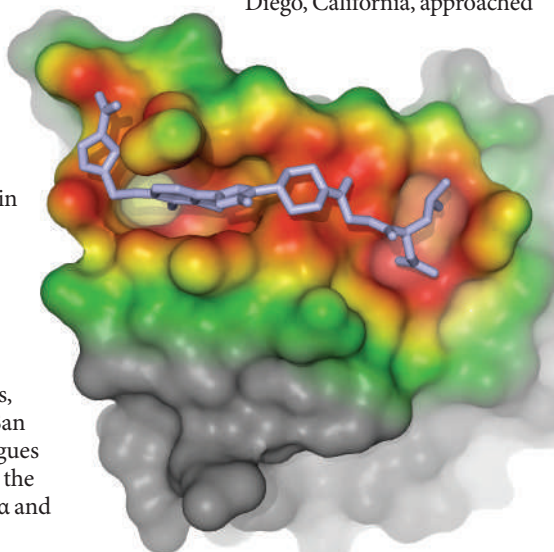
GEOLOGY

Deep impact

Earth Planet. Sci. Lett. doi:10.1016/j.epsl.2006.09.009 (2006)

A painstaking survey of rocks from around the globe has provided new information about the nature of a meteorite impact 65 million years ago, which may have triggered the mass extinction that wiped out the dinosaurs.

The impact would have been most devastating if the meteorite hit at a shallow



DELANO SCIENTIFIC LLC

angle, digging through the Earth's top layers. This could show up as asymmetries in the distribution of ejecta around the crater site in Chicxulub, Mexico.

Joanna Morgan of Imperial College London, UK, and her colleagues examined nearly 5,500 shocked quartz grains from sites as far afield as Italy and New Zealand. At most, the researchers report, there is evidence that ejecta lying to the southeast of the crater experienced greater shock, which, along with thick ejecta deposits in Belize, would be consistent with an impact angle of 45° or more.

CLIMATE RESEARCH

Winds of change

J. Clim. **19**, 5388–5404 (2006)

Winds that blow over the mountains of the Antarctic Peninsula are the main driving force behind the exceptional warming in this region, report Gareth Marshall of the British Antarctic Survey in Cambridge, UK, and his colleagues. That warming has contributed to the collapse of ice shelves surrounding the peninsula.

Marshall's team used data from climate stations to reconstruct wind patterns over the past 50 years. They found that the strength of the westerly circumpolar wind had increased over this period, which is an expected effect of global warming and ozone depletion in the Southern Hemisphere. The stronger winds more frequently made it over the 2,000-metre-high mountains that divide the peninsula. Air warms as it descends, creating 'föhn' winds that raise local temperatures.

STEM CELLS

Caution urged for therapy

Nature Med. doi:10.1038/nm1495 (2006)

Brain cells that produce the neurotransmitter dopamine have been created from human embryonic stem cells and used to treat rats with a model of Parkinson's disease. But a potentially cancerous side effect might put the brakes on developing such therapies for humans.

It's currently difficult to turn stem cells into the right kind of neuron in large enough numbers for useful therapies. Steve Goldman of Cornell University Medical College in New York and his colleagues encouraged stem cells to grow into dopamine-producing neurons by placing them in cultures similar to the environment of the developing brain. Transplanting the neurons into the rats restored much of the animals' lost movement. But an expanding core of undifferentiated

stem cells within the mass of transplanted dopamine cells could cause tumours, the researchers warn.

ANIMAL BEHAVIOUR

Extreme diving

J. Exp. Biol. **209**, 4238–4253 (2006)

Scientists monitoring two species of beaked whale (*Mesoplodon densirostris*, pictured below, and *Ziphius cavirostris*) have recorded ocean dives that are among the deepest and longest reported for air-breathing species.

The dive data came from devices attached



M. USHIODA/IMAGEQUESTMARINE.COM

by suction cups to 10 whales, which were recorded making 44 dives near Italy and the Canary Islands. Some individuals spent nearly an hour at depths of up to 1,888 metres. The whales also engaged in a series of shallower dives after each deep prey-foraging trip.

Lead author Peter Tyack of the Woods Hole Oceanographic Institution, Massachusetts, speculates that beaked whales' ability to dive so deep has implications for understanding how the mammals may be affected by sound from devices such as military sonar.

OPTICS

Keep it together

Nature Phys. doi:10.1038/nphys438 (2006)

The ability to slow down pulses of light could be useful in telecommunications. But slowed pulses have a bad habit of spreading out, degrading the information that they carry. Joe Mok and his co-workers at the University of Sydney, Australia, present a cure.

The researchers modified the properties of an optical fibre in a repeating pattern along its length. They showed that sub-nanosecond pulses of light travelled through this 'fibre Bragg grating' at 16% of the typical speed of light. What's more, each pulse behaved as a soliton, a type of wave that shows no dispersion. This happens because the pattern reflects the weak edges of a spreading pulse back towards its brighter centre.

JOURNAL CLUB

Shigeru Kuratani
RIKEN Center for Developmental
Biology, Kobe, Japan

An evolutionary biologist hopes to tell head from trunk.

Anatomy involves dissecting the body not only into its physical components, but into categories and concepts as well. This is indeed the only means by which we can bring an approximation of logical order to our understanding of biological phenomena, or to describe and explain them to others.

Unfortunately, the results of anatomical parsing tend to be semantic. Just as a single word can convey different meanings depending on its grammatical context, so can anatomical names apply to different parts of an embryo during development.

The boundary between 'head' and 'trunk', for instance, has resisted clarification — I use at least five different definitions according to the embryological events on which I happen to be focusing.

An article published last year appeared to only add to the menu of choices by proposing that an intermediate 'neck' separates the head and trunk (T. Matsuoka *et al.* *Nature* **436**, 347–355; 2005).

The authors looked at a cell population known as the neural crest. This runs the length of the embryo and forms a range of structures. They defined the neck as the region where crest cells that differentiate as they do in the head overlap with a second structure that is characteristic of the trunk. That structure consists of segments of tissue known as somites.

Naming this grey area a 'neck' in fact neatly sidesteps the need for a clear demarcation between head and trunk. It allows the anatomical terms to be flexible at their boundaries.

But, sadly, it hasn't helped me. In my lab we study lampreys, a kind of ancient fish, and its embryos don't seem to have a neck. I suppose biological concepts also have their evolutionary origins, and the lamprey must pre-date the 'neck'.

SPECIAL REPORT

'A shocking lack of evidence'

The trial in Libya of six medics accused of infecting children with HIV ends next week. With the defendants facing possible execution, **Declan Butler** asked AIDS experts to assess the case against them.

A scientific report being used in the case against six foreign medical workers facing the death penalty in Libya is nothing but conjecture and supposition, say international experts. The evidence, commissioned from Libyan medics, has bolstered the charge that the six knowingly infected more than 400 children with HIV at the Al-Fateh Hospital in Benghazi in 1998.

The court has denied requests by defence lawyers to hear evidence from international experts. Instead, five Libyan physicians testified in August that they stand by the conclusions of their 2003 report, commissioned by the court in an earlier trial, and on 29 August the prosecution called for the medics to be given the death penalty. The trial is due to end on 31 October (see 'Countdown to a verdict').

With a guilty verdict looking likely, *Nature* obtained an English translation of the Libyan report, which has been key to the prosecution's case, and asked leading international experts to assess it.

"I don't see any evidence in it," says Janine Jagger, an epidemiologist and MacArthur fellow who heads the International Health Care Worker Safety Center at the University of Virginia in Charlottesville. "It wouldn't meet the

lowest standards of epidemiological evidence for establishing any causal relationship."

In 2003, the court also ordered a report from Luc Montagnier, who discovered the AIDS virus and is president of the World Foundation for AIDS Research and Prevention, and Vittorio Colizzi, an AIDS researcher at Rome's Tor Vergata University. They concluded that the infections were caused by poor hospital hygiene, and started before the medics arrived in Libya (see "Montagnier and Colizzi's conclusions"). But the court threw out this report, on the grounds that the Libyan panel had reached the opposite conclusion. The panel had dismissed the external report as "hypothetical" and "lacking precision".

'That's tosh'

"The [Libyan] report refutes the Montagnier and Colizzi report on the grounds that there is no written record of the reuse of injecting equipment, and a blank denial that indwelling catheters were ever used," says Robin Weiss, an AIDS virologist at University College London. "It wrongly turns lack of evidence into evidence of absence."

The report argues that HIV and hospital



hygiene were not a problem in Libya (the prosecution describes the Al-Fateh Hospital as a "model") and that the outbreak is so large that deliberate, malicious infection of HIV cannot be excluded. "I don't agree with that statement," says Weiss. "And even if I did, it does not amount to sufficient evidence to incriminate the accused medical staff."

The Libyan report also suggests that because the genetic sequence of the Benghazi HIV strain is different from any lodged in public databases, there are grounds for suspecting foul play. "That's tosh," says Weiss. Montagnier agrees, pointing out that the virus was a new natural recombinant of a highly infectious

Countdown to a verdict

The next, and last, session of the trial of five Bulgarian nurses and a Palestinian doctor, on charges that could see them face the death penalty (see main piece), is scheduled for 31 October, with a verdict expected days or weeks later.

The retrial began on 11 May, with a handful of sessions since, but little new evidence. The medics were condemned to death in May 2004, but the Supreme Court quashed their convictions last December, following international protests.

International pressure eased

once a retrial was won, but Emmanuel Altit, head of the defence team, is pessimistic about the upcoming verdict, from the same Benghazi Criminal Court. "The decision will more than likely be a very bad one," he predicts.

As the trial draws to a close, many scientific and human-rights

The six medical workers' retrial has been running since May.



bodies have renewed calls for the court to hear independent evidence (see *Nature* **443**, 612–613; 2006). The New York Academy of Sciences and the Federation of the European Academies of Medicine have recently launched campaigns, and this week, *Science* will publish a letter calling for the release of

the six, signed by AIDS researcher Robert Gallo and dozens of other scientists.

"It's exactly what we need, to pressure the Libyan authorities," says Altit. Even if too late for the current trial, such pressure might help influence the Supreme Court, who would hear an appeal if the medics are found guilty.

Altit believes the medics were charged so that officials wouldn't have to admit responsibility for the healthcare system's failings. But the case is embarrassing for the government, he says, so there is hope of a resolution.

D.B.

M. TURKIA/AFP/GETTY



WEB FOCUS

AIDS medics in Libya: all our coverage, plus blogs, links and key documents.
www.nature.com/nature/focus/aidsmedicslibya

S. NENOV/REUTERS



T. PENOV/BTA/REUTERS

Bulgarian president Georgi Parvanov (left) visits HIV-positive children in Benghazi, Libya, in 2005.

strain common in Central and West Africa, which has replaced most other strains in the region over the past few years.

In contrast, Weiss describes Montagnier and Colizzi's report as excellent. "Colizzi has done a really superb job in difficult circumstances," he says. After studying both reports, Weiss concludes: "There are no grounds for suspicion of deliberate infection by any staff, and strong evidence of hospital-acquired infection before the arrival, and after the departure, of the Palestinian physician and the Bulgarian nurses."

'Completely inadequate'

Jagger, an expert in occupational exposures to blood-borne pathogens, says she is astonished that the medics were even arrested, given the flimsiness of the prosecution's scientific evidence. In this sort of case, she says the minimum standard should be a thorough field study that tracks all medical procedures carried out during the outbreak, and calculates attack rates, epidemic curves and other standard epidemiology measures for inferring cause. She describes the Libyan data as "completely inadequate".

To firmly establish any cause, a case-control study should also have been done, Jagger adds, comparing risk factors and medical procedures used between the HIV-infected patients and a similar uninfected group to try and explain transmission. "Inexcusably, there has been

Montagnier and Colizzi's conclusions

- Many of the children were also infected with hepatitis B and C, suggesting unsafe medical practices were common in the hospital.
- Two nurses at the hospital were infected with the same strain of virus as the children, again suggesting poor hygiene.
- The outbreak began by 1997, before the accused medical workers arrived in Libya.
- The virus responsible belongs to the HIV-1 subtype A/G, a recombinant strain common in Central and West Africa that is highly transmissible and virulent — not a genetically modified strain as suggested by the Libyan authorities.
- There is no evidence for deliberate injection of HIV-contaminated material.

no attempt by Libyan officials to conduct an epidemiological study that could establish a causal link between the infected patients and individual care givers."

Luc Perrin, a clinical virologist at Geneva University Hospital in Switzerland, who has treated many of the infected children, describes the Libyan report as "a lot of generalities that are not always correct". The report also fails to provide any evidence for its assertion that HIV infection has not been seen in children at other Libyan hospitals, he says.

Perrin is an expert on primary HIV infection. He has analysed samples from 148 of the infected children, collected in September 1998, and has obtained further data on 37 of them and 46 of their parents, when they were treated in Switzerland. Perrin says his genetic data support Colizzi's analysis, and that many of the 1998 samples have protein profiles corresponding to infections well over a year old: "I can tell for sure that the HIV infection cases occurred before September 1997 and the first cases most likely before 1996." The accused medics first arrived in Libya in March 1998.

The Libyan report is also silent on the prevalence of hepatitis at Al-Fateh Hospital and other Libyan hospitals, notes Perrin — who found that half of the HIV-infected children were also infected with hepatitis B or C. He says these high levels "clearly indicate" that the children were exposed to infection via contaminated blood or other medical material. Moreover, many of the children were infected with several subtypes of hepatitis, suggesting they were exposed to hospital contamination on multiple

occasions, possibly when receiving vaccination injections. "If a single source of contaminated blood had caused the HIV outbreak, all the children would be infected by the same hepatitis C subtype," says Perrin. "What we observed can [instead] be explained by the reuse of syringes or poor sterilization procedures."

Perrin believes the most likely scenario is that a child who was infected with HIV in Al-Fateh in 1997 or earlier, returned to the hospital in 1998. "The child now is highly infectious, so poor medical procedures or sterilization procedures will rapidly translate into a number of new HIV infections."

'Sinister legacy'

Perrin was commissioned to write reports for the Libyan government in 2000 and 2001, and for President Muammar Gaddafi in 2004, but says that they received no response. "It is strange that I was asked by the World Health Organization representative of Libya to investigate, and received a grant for that, that I sent the reports accordingly and finally the report is not considered," he says. Testimony he submitted to the court was also rejected, he adds.

The purported 'smoking gun' in the Libyan report is the detection of HIV antibodies in vials allegedly found at one of the nurses' homes during a raid in 1999, but not tested until 2003. Both Montagnier and Colizzi have seen the results of a western blot, a test to detect proteins: they are "indeterminate", says Montagnier. "They say nothing," adds Colizzi. In 2002 Libya promised that they could test the samples independently, but neither has ever been given access.

Even a positive test could detect only antibodies to HIV. It would not show that the vials had contained the virus, points out Massimo Amicosante, a biologist also at Tor Vergata. "This is one of the main weak and controversial points of the Libyan report," he says. Finding the virus would require testing for HIV RNA, which has not been done.

After reviewing the evidence, experts are in no doubt as to the consequences of a guilty verdict. "If the accused are found guilty, it will be a travesty of justice," says Weiss. "Moreover, it would be foolhardy for any expatriate healthcare worker, whether from the Arab world or elsewhere, to work in Libya."

Jagger's conclusions are similar: "There is a shocking lack of evidence in this case," she says. "The Libyan government stands to carry out an act that would not be forgotten by the international healthcare community. Such a horrific humanitarian tragedy would stand for ever as a sinister legacy of the Libyan government." ■

"There are no grounds for suspicion of deliberate infection, and strong evidence of hospital-acquired infection."

ZOO NEWS

The not very hungry caterpillar

A newly discovered caterpillar may be the world's thinnest. The larval stage of the moth *Houdinia flexilissima* lives inside the stems of cane rushes. At a few centimetres long but less than a millimetre wide, its discoverers have named it 'Fred the Thread'.

SCORECARD

**Solar energy**

The Internet company Google is planning a solar electricity system for its headquarters in Mountain View, California. With a capacity of 1.6 megawatts, the installation will be one of the largest on any corporate site.

**Celebrity genomes**

Genomics entrepreneur Craig Venter famously raced the Human Genome Project to sequence the human genome. He has now almost finished a map of his own genome — the first of any individual — and plans to make it public.

ON THE RECORD

“I doubt there's a scientist in the country who fits the description.”

As Kazakhstan is promised a 25-fold increase in science spending, Tynysbek Kalmenov, director the Centre for Physics and Mathematics in Almaty, is pessimistic about finding world-class researchers to head the resulting institutions.

“A light bulb goes on in my head, and energizes my hand to begin to sketch out some scribbles.”

Architect Frank Gehry explains the creative process to an audience of neuroscientists.

“Experts: some women perform well in math.”

The Associated Press has trouble summing up a study on the influence of cultural factors on women's mathematics ability.

Sources: Australian Broadcasting Corporation, Electronics Weekly, Sunday Times, IWP.



K-PHOTOS/ALAMY

Sceptics detect flaws in US nuclear monitor plan

WASHINGTON DC

Now that North Korea has joined the nuclear club, the United States is intent on preventing it from extending membership perks to others.

Secretary of State Condoleezza Rice visited the region last week to meet leaders from neighbouring countries and discuss ways to detect and intercept illicit nuclear stocks. But proliferation experts disagree about whether such a screening regime is practical or even possible. Spotting radioactive material aboard ships, trucks and aircraft is technically difficult and would require unprecedented regional cooperation.

North Korea's nuclear test on 9 October (see *Nature* 443, 610–611; 2006) renewed fears about nuclear proliferation. Last week, Mohamed ElBaradei, head of the United Nations' nuclear watchdog, the International Atomic Energy Agency, warned that another 20 to 30 states could soon have the capacity to develop nuclear weapons.

A particular concern is that North Korea might smuggle nuclear material to other states or terrorist groups, says Michael Burns, principal deputy for threat reduction at Los Alamos National Laboratory in New Mexico. Burns says that before leaving for her trip, Rice's team consulted Los Alamos on the possibility of detecting nuclear materials leaving North Korea.

Detection is possible, says Burns, but only by integrating a lot of information: “Detection of nuclear material needs a systems-wide approach.” The most important tactic will be to create bottlenecks for goods travelling in and out of North Korea. These ports and border crossings will then need a variety of detection equipment. Passive detectors catch flashes of gamma radiation from the decay products of uranium or plutonium, or neutrons from the decay of certain radioactive isotopes, notably plutonium-240. In addition, active detectors use X-rays or neutrons to scan shipping containers and trucks for dense masses. “That will be an immediate indication to a customs agent to take it aside

and look at it by eye,” says Burns.

If nuclear material were found entering or leaving the country, it would also be important to establish where it came from — which is difficult without international cooperation, says Burns: “We have to establish information-sharing through diplomatic channels.” One proposal is to create an international database of nuclear samples (see description on page 907). It's hoped that the ability to trace the origin of illicit stocks would also deter states from selling their nuclear material.

But others disagree about whether such a detection scheme could work. “It's not feasible,”

“This is not something you can do from a mile away. You have to get as close as possible.”

Polish scientists fight creationism

Fifty leading scientists in Poland have signed an open letter in protest against an aggressive anti-evolution campaign launched by the League of Polish Families (LPR), the ultra-right-wing coalition partner in the conservative Polish government.

“The theory of evolution is a lie,” Mirosław Orzechowski, Poland's deputy education minister, told the newspaper *Gazeta*

Wyborcza on 14 October. “It is an error we have legalized as a common truth.”

The LPR entered the ruling coalition in May 2006. Its leader, Roman Giertych, is also known to favour creationist views. These, as well as his openly homophobic, anti-Semitic and nationalistic opinions, have sparked student demonstrations in Warsaw since he took the minister of education job in May.

Giertych's father, Maciej Giertych, is an LPR member in the European Parliament and is lobbying for obligatory inclusion of creationism in Polish biology curricula. Maciej, who holds a PhD in tree physiology from the University of Toronto, Canada, claims darwinian evolution is refuted by scientific evidence.

Orzechowski's comments have rattled Poland's science



K. SASAHARA/AP

Shipping news: plan to step up detection amid fears that North Korea may smuggle nuclear material.

says Steve Fetter, dean of public policy at the University of Maryland in College Park. “The problem is that the radiation signature from the nuclear materials you use in weapons is quite small.” It would be relatively easy to shield the radiation with lead, plastic and other materials, he says. Gaining access to North Korean cargo is also complicated. “This is not something you can do from a mile away,” Fetter says. “You have to get as close as possible.”

Daniel Pinkston, director of the East Asia non-proliferation programme at the Monterey Institute of International Studies in California, agrees, adding that borders are porous and difficult to monitor. “And boarding ships on the high seas is basically an act of war.” Beyond the

technical difficulties, Pinkston thinks North Korea would be unlikely to surrender its limited plutonium stocks. “Why would North Korea give up its fissile material? It’s priceless to them,” he says. A more real concern, he believes, is that North Korea may export nuclear knowledge: “Data sets or design plans can’t be interdicted at all.”

Burns declined to comment on the ability to detect nuclear material remotely, but says he believes a perimeter around North Korea can be built. “Nothing is 100% effective,” he says. “But a system that increases our confidence that we are deterring the movement of nuclear material can be established.” ■

Geoff Brumfiel

community. Researchers are concerned that the LPR campaign could infiltrate biology teaching in schools.

Maciej Żylicz, a senior researcher at the International Institute of Molecular and Cell Biology in Warsaw, says he was “shocked” by the remarks. “We really did not expect a creationist movement to emerge in Poland.”

“It is a catastrophe,” adds Bartosz Borczyk, who is completing a PhD in zoology at the University of Wrocław and wrote to

Nature about the issue.

“People could easily get the impression that there is a controversy about evolution among scientists.”

Michał Seweryński, Poland’s minister of science, has criticized the LPR’s position. “There is no need for a discussion,” he told *Nature*. “Scientific evidence is clear and the opinion of a minority will not change teaching in schools.”

Members of the Polish Academy of Sciences protested against the LPR campaign in an open letter

that was published in several Polish newspapers on 17 and 18 October. Żylicz, who signed the letter, says he hopes the quick response will avert damage to Polish science and education. “However, the point that really requires further discussion is not evolution, but how a minister can say such stupid things.”

Neither Roman nor Maciej Giertych, nor Orzechowski, responded to *Nature*’s request for comment. ■

Almut Graebisch



RACE TO SPACE IN NEW MEXICO

Find reports of the annual space-elevator competition online.

www.nature.com/news

SPACEELEVATORBLOG.COM

From hive minds to humans

Could an insect help us understand why some people are daredevils, or why others are overweight? Scientists who have sequenced and analysed the genome of the western honeybee *Apis mellifera* say it might. The sequence, released in this week's *Nature* (see page 931), is already aiding the young field of sociogenomics — the search for the genetic underpinnings of social life.

Researchers have shown some examples of how genes, the environment and behaviour can influence — and be influenced by — social interactions. Implanting one gene, for example, can change voles from promiscuous to monogamous (M. M. Lim *et al. Nature* **429**, 754–757; 2004). And maternal care from rats helps pups to express genes that make them less susceptible to stress as adults (I. C. G. Weaver *et al. Nature Neurosci.* **7**, 847–854; 2004).

But, besides its economic importance (see 'Pollinators in peril'), the honeybee is the lodestone of social biology. "Everything they do is social," says Charles Whitfield, an entomologist and molecular biologist at the University of Illinois at Urbana-Champaign. The colony dynamic determines a bee's diet, work and sex life, making bees ideal for studying some aspects of social biology, say scientists.

Researchers are also releasing a tool set for analysing the genome, including an improved microarray of honeybee genes and a list of DNA variations called single nucleotide polymorphisms. Scientists hope to use these tools to find the genetic causes of social traits — such as the bees' dance language and the division of labour in the hive. "We have so many behaviours in the bee for which we don't have any clue about the mechanism," Whitfield says.

The genome is helping to reveal some of those mechanisms. For instance, there are 65 spots in the genome that seem to code for short



Work programme: a bee's job is determined by how its genes interact with its environment.

RNA molecules called microRNAs (miRNAs), molecular switches that can turn genes on or off. The researchers found that miRNA activity differs between bees doing different jobs. Such work starts to show how gene regulation gives the bee such a different lifestyle from, for example, the fruitfly, with which it shares many genes. "At the level of the brain, the parts list is pretty much the same as the fruitfly, yet there are these radical differences in behaviour and social organization," says George Weinstock, a genome biologist at Baylor College of Medicine in Houston, Texas. "That makes you think that regulation is probably more important than the parts list."

So, what does this have to do with people? Primate behaviour specialists say that there are parallels between honeybee and human socie-

ties. For instance, some individuals in each care for youngsters who aren't their offspring. But, cautions anthropologist Sarah Blaffer Hrdy of the University of California, Davis, honeybees and humans have many differences. "Anyone

"We have so many behaviours for which we have no clue about the mechanism."

interested in the cognitive or emotional bases of cooperation should stick with other primates, or at the very least other social mammals," Hrdy says. "The relevance of honeybee genomes would depend very

much on the sort of questions one is asking."

Bee researchers say there could be links between human and bee behaviour. For instance, some bees seek out food sources, others wait and follow on. It's possible that scouts have the same sort of novelty-seeking genetic pathways that predispose some humans to become bungee jumpers. "One of the lessons of developmental biology is just how conserved gene function is in building complex phenotypes" such as social behaviours, says Gene Robinson at the University of Illinois at Urbana-Champaign.

Some even see a link to human health. Worker bees make queens by feeding larvae a protein-rich substance called royal jelly. Understanding how the jelly reprogrammes genes might tell us more about how humans get fat, says molecular biologist Ryszard Maleszka at the Australian National University in Canberra. "Some of the insights we gain from following this protein family might be useful for probing a major problem in human society: obesity." ■

Erika Check

See Editorial, page 884; News & Views, page 919.

Pollinators in peril

Honeybees aren't just busy socialites. They're also crucial for agriculture, pollinating 90 major commercial crops in North America alone. But the honeybee and other pollinators are in danger. In a report issued on 18 October, the US National Academy of Sciences (NAS) found that pollinators such as bees, birds and bats are declining

rapidly in North America. The report recommends that North Americans set up better monitoring systems for pollinators. Such systems are already in place in Europe, where scientists have documented the extinction of some pollinating species.

The agricultural bee-keeping industry will probably ensure the honeybee's

survival, but populations have still dropped by 30% over the past 20 years in North America, mainly due to the varroa mite. Scientists are working on breeding strains that are resistant to the mite, although the NAS report calls on industry to do more on this front. The genome sequence should help to speed the process up.

E.C.

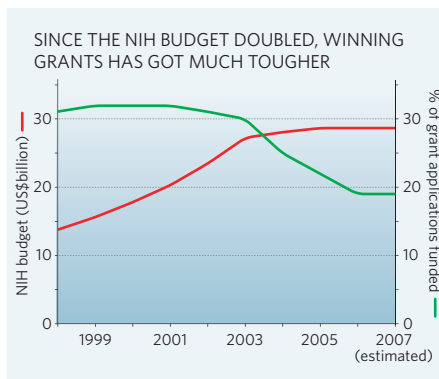
Grants fall victim to NIH success

Seven thousand fewer people showed up at this year's Society for Neuroscience meeting than last year. That still left 21,000 researchers attending one of the most popular conferences in biomedicine. But hallway gossip swirled around one possible cause for the drop-off — money, in the form of grants from the US National Institutes of Health (NIH).

"Dollars come to mind," says Jeffrey Rothstein, a neurologist at Johns Hopkins University in Baltimore, Maryland. "The field has expanded and we now have less money per person."

Rothstein's comments explain an apparent paradox: the annual budget for the NIH doubled between 1998 and 2003, and now stands at \$28.6 billion. But many researchers are finding it far more difficult to win grants. This is because universities invested hugely in infrastructure and workforce once the doubling began, which led to an explosive growth in grant applications. In 1998, the NIH received 24,000 grant applications; this year the number is 46,000 and is forecast to reach 49,000 in 2007.

Two other metrics have also changed, helping to create the perfect storm. Average grant sizes have grown by 40% since the doubling began, from \$275,000 in 1998 to \$400,000 today; and Congress has stopped increasing the NIH budget. So the agency's buying power



has fallen in real terms by 5.5% since the doubling ended in 2004.

"Nobody's bottom-line model showed funding decreasing after the doubling. We were talking about modest increases, not the floor falling out," says Tony Mazzaschi of the Association of American Medical Colleges (AAMC) in Washington DC.

In a speech at the neuroscience meeting, NIH institute heads acknowledged the inevitable outcome: the success rate for grant applications across the agency dropped from 31% in 1998 to 23% last year (see graph). Next year, this is expected to fall to 19%. "We have really good grants that we can't afford to pay," says Story Landis, head of the National Institute of Neurological Disorders and Stroke.

"The demand took off just as the NIH budget was landing."

With Congress highly unlikely to hand the NIH another significant increase in the near future, researchers and agency officials are considering how to adjust the funding system. Some are dubious that much pain can be avoided. "I personally don't see a terribly satisfying solution," says David Korn, former dean of the Stanford University School of Medicine in California and a top official at the AAMC who has studied the issue extensively.

Elias Zerhouni, the NIH director, is battling to counter that notion with a number of short-term strategies, especially for first-time applicants and those fighting to get expiring grants renewed. "My main concern is the individual investigator whose lab might close because of the gap in funding," he says. This year, Zerhouni cut all existing grants by 2.35%, to free up money for new grants and grants competing for renewal. He will also use funds from expiring grants to gain a 3% boost in the number of new and competing grants awarded next year.

For some, these measures will be too little, too late. Rothstein says that he may have to lose staff and is already seeing "pruning" in other labs. His brother Jay, a cancer researcher at Thomas Jefferson University in Philadelphia,

SOURCE: NIH

Funding agencies toughen stance on open access

The Howard Hughes Medical Institute (HHMI), a non-profit research organization that funds more than 300 US researchers, is considering a plan to pressure its investigators into making their published papers freely accessible.

The plan, if approved, would dictate that publications must be deposited in a public database within six months of publication in order to count towards an investigator's application for reappointment. HHMI investigators apply for reappointment every five or seven years.

The proposal embodies the latest stage of the open-access movement: enforcement. After

years of requesting voluntary compliance, several funding agencies are considering tougher stances. In 2005, the US National Institutes of Health (NIH) asked grantees to voluntarily deposit articles in a public database such as PubMed Central within 12 months of publication. A year after the request, only 4% of NIH grantees had done so, prompting Congress to propose legislation mandating compliance.

Meanwhile, on 1 October, Britain's Medical Research Council and the Wellcome Trust medical charity began requiring grantees to deposit final, peer-reviewed manuscripts in public databases as soon as possible, but no later

than within six months of publication. Failure to do so, says Wellcome Trust director Mark Walport, would be a breach of granting conditions.

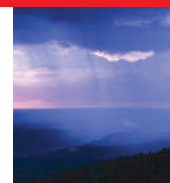
HHMI president Thomas Cech says a decision on the proposed policy will probably be made in early 2007. Although HHMI officials say they will not legislate where their investigators publish, several researchers say the threat of weakening their reappointment application represents significant pressure.

Cech says the proposed policy is simply an extension of HHMI guidelines about sharing published reagents and other research material. "The publications are

the most useful product of our investigators' research," says Cech.

The HHMI is still negotiating with publishers, but Walport says most major journals, including *Science* and *Nature*, have complied with the Wellcome Trust's guidelines.

Many publishers let authors pay to make their articles available immediately. For example, Wellcome Trust grantees can make their papers in most Elsevier journals publicly accessible for \$3,000 per article. Both the HHMI and the Wellcome Trust already provide funds for publication in open-access journals. "We see payment to the publisher as part of the cost of research," says Walport. "And we're prepared to pay appropriately."



RAIN MAKES THE GROUND SHAKE

A wet weekend may be enough to set off an earthquake.

www.nature.com/news

Pennsylvania, will quit for the pharmaceutical company Amgen early next year, in part because of the struggle to get funding. He notes that few graduate students at his university plan to remain in academia: "People see how hard it is to run a lab."

After the NIH presentation at the neuroscience meeting, Peter Cariani spoke up. He studied spinal-cord regeneration at Tufts Medical School in Boston, Massachusetts, until his lab ran out of funding two months ago. He won applause for suggesting that the NIH should "make it harder for the rich to get progressively richer". Like others present, he feels that big labs Hoover up too many grants, forcing teams like his out of the game. He wants the criteria for second and subsequent grants to be set higher than for an initial application.

Zerhouni cautions that such proposals put key research decisions into the wrong hands. "It would be very dangerous to dictate by bureaucratic fiat that somebody who is meritorious should not be funded because they have a grant or two. The peer reviewers are smart enough to know if someone is being overfunded or not."

He hopes scientists will stick out the hard times: "We still have the largest research budget on Earth: \$28.6 billion is larger than all other countries combined. If you believe in your research, don't give up on it. Things are not great. But they are not as desperate as some people portray them to be."

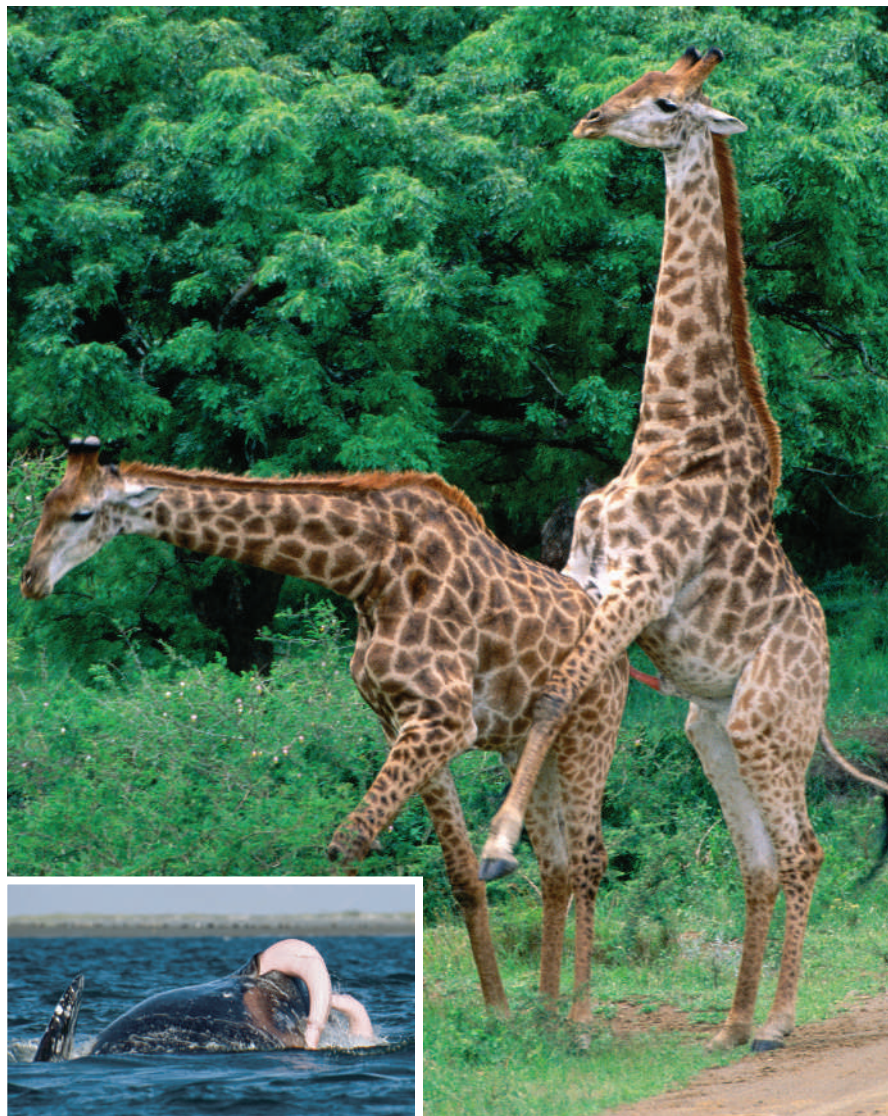
Jim Giles and Meredith Wadman

But some publishers continue to struggle with the Wellcome Trust's requirement. If a paper is available in a public database, this reduces the number of visits to subscription sites, says Martin Frank, executive director of the American Physiological Society.

Frank says this is why his society refuses to comply with the Wellcome Trust's guidelines. But such decisions may be more difficult with the proposed expansion of open-access requirements — Wellcome Trust-funded research accounts for only 2% of the papers in the society's 14 journals, but the NIH supports almost half.

Kathleen Case, publisher of the American Association of Cancer Research's five peer-reviewed journals, also considers the requirements too strict. "The Wellcome Trust asked us to change our policy and we said no," says Case. "And now we've been blacklisted."

Heidi Ledford



SNAPSHOT But is it natural?

When Norway's museums authority called for ideas for an exhibition that engaged with contemporary societal debates, Geir Söli, head of exhibitions at Oslo's Natural History Museum, had just the theme. He had been listening to a priest arguing on the radio that homosexuality is a sin and contrary to nature.

People can only make up their minds about whether such behaviour is 'unnatural' if they know the current state of research on homosexuality in nature, he reasoned.

No sooner had the authority approved his proposal than church groups began to protest against this use of public money. But *Against Nature? An Exhibition on Animal Homosexuality* opened peacefully at the Natural History

Museum on 12 October to wide acclaim. It will run until next August.

Söli focused the exhibition around 50 species, out of the 500 or so in which homosexual behaviour has been well documented. The exhibition comprises photographs, stuffed animals, models and other objects, along with explanatory text in Norwegian and English.

Two male giraffes are shown here indulging in roadside sex. Such encounters have been recorded in the literature as involving anal penetration and ejaculation. The inset shows two male whales (*Eubalaena australis*) engaged in sexual games.

Alison Abbott

A. VAN ZANDBERGEN/LPI

S. WONG/MARINETHEMES.COM

Embattled IBM paper published after court row

It took a fight in court, but a controversial paper that links working in an IBM factory to increased rates of cancer has finally been published.

The work analyses nearly 32,000 people who had been employed by IBM, and suggests that workers died from certain cancers at significantly higher rates than seen in the general US population (R. W. Clapp *Environ. Health* doi:10.1186/1476-069X-5-30; 2006).

Two years ago, the journal *Clinics in Occupational and Environmental Medicine* refused to publish the work, saying it was unsuited to a review journal. Journal contributors alleged that the publisher had succumbed to pressure from IBM, an allegation the publisher denied (see *Nature* 429, 687; 2004).

The data were released as part of a lawsuit against IBM, in which the author — Richard Clapp, a professor of environmental health at the Boston University School of Public Health in Massachusetts — acted as an expert witness. IBM has argued that the data were methodologically flawed and released for use only in the court case. Clapp says a New York court gave him the go-ahead this spring to publish.

Iceland breaks ban on commercial whaling

Icelandic fisheries officials announced on 17 October that they intend to resume commercial whaling, allowing the catch of 30 minke and 9 fin whales by the end of August 2007. Within days, whalers went ahead and took their first kill: an endangered fin whale.

The International Whaling Commission, which has had a moratorium on commercial whaling since 1986, estimates that there are some 25,000 fin whales in central North Atlantic waters, and more than 40,000 minke in Iceland's coastal waters. Iceland's ministry of fisheries argues that the proposed catch is

sustainable. But conservationists and other governments have criticized the move.

Norway is the only other nation that allows commercial whaling, taking some 600 minke per year. Inuit hunters from Greenland and Canada also take a small number of whales, and Japan and Iceland undertake controversial scientific hunts.

Physicist retires to work on 9/11 conspiracy theories

A physics professor known for his theories that bombs, not planes, destroyed the World Trade Center on 11 September 2001 is retiring from his position at Brigham Young University in Utah.

University officials had placed Steven Jones on paid leave last month while reviewing his work on 9/11 conspiracy theories. On 20 October, while the review was still under way, the university announced that Jones had reached an agreement to take early retirement. Jones says he plans to continue his 9/11 research.

The researcher supports his theories with analyses of soil and metals retrieved from the collapsed World Trade Center, but the university's faculty and others have questioned the rigour of his research. The American Association of University Professors questioned the university's decision to place Jones on leave while reviewing his work, saying it infringed his academic freedom.

Agent of Singapore's drive to biomedicine steps down

Can Singapore's leading research organization survive the departure of its charismatic leader? Philip Yeo, the chairman of Singapore's Agency for Science, Technology and Research (A*STAR), has announced that he will be stepping down in March 2007. He is to become chairman of the Standards, Productivity and Innovation Board, and a science adviser to the minister for trade and industry.



The thriving Biopolis science complex in Singapore is to get a new chairman.

Yeo has been the driving force behind the creation of the city-state's thriving Biopolis complex, a collection of bioscience institutes that has drawn international talent (see *Nature* 436, 767–768; 2005). David Lane, who heads the Institute of Molecular and Cell Biology within A*STAR, says that Biopolis is now well enough established to attract top names without Yeo. "We seem to have passed some barrier of critical mass now as we have loads of unsolicited applications," he says.

Yeo's successor, Lim Chuan Poh, who has held science-policy-related positions within A*STAR and at the education ministry, is popular with Yeo fans and critics alike. One of the latter says Poh seems to be "inclusive, progressive and impressive — an excellent choice".

European weather satellite reaches polar orbit

Europe has launched the first of three polar-orbiting weather satellites, which will eventually be joined by a long-troubled US satellite system.

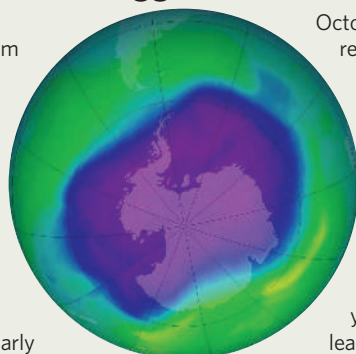
After several delays, the MetOp-A satellite was launched from the Baikonur launch facility in Kazakhstan on 19 October. It now sweeps around the globe from pole to pole 14 times a day, and should soon start transmitting data from a variety of instruments — including a water-vapour profiler and a spectrometer to measure trace gases such as ozone.

It joins a cluster of polar-orbiting satellites operated by the US National Oceanic and Atmospheric Administration. These are meant to be replaced by an \$11-billion system jointly operated with NASA and the US military, but delays and cost overruns have threatened those plans in recent years (see *Nature* 441, 798; 2006).

Antarctic ozone hole is bigger than ever

As ozone-depleting chlorofluorocarbons fade from the atmosphere, the ozone hole over the Antarctic continues to fluctuate in size. And this year, it was the biggest it has ever been.

Both NASA and the European Space Agency documented the size of the 2006 ozone hole and found that, in late September and early



October, it had broken records for both size and depth, previously set in 2000 and 1998. This image, taken on 24 September, shows the hole at an extent of 29 million square kilometres. Temperatures above the Antarctic have been the lowest this year since 1979, which helped lead to the record ozone loss.

BUSINESS

Capital gains

London is seeking to compete with the Nasdaq as the preferred global venue for science-based companies trying to raise cash. **Andrea Chipman** reports.

When James Karis needed to raise funds last summer for his company Entelos, based in Foster City, California, he ended up looking far afield. Entelos, which builds computer models of disease for clients in the drug industry, raised the money through an initial public offering on London's Alternative Investment Market (AIM).

For young companies around the world — including US ones wary of the increased costs of listing on the Nasdaq — it's an increasingly popular route. Clipper Windpower, a wind-turbine designer based in Carpinteria, California, and Elcom, a software company based in Norwood, Massachusetts, are among some four dozen US firms to list on the AIM since 2004.

But the approach has pitfalls. Relative to the Nasdaq, observers say, the AIM still lacks specialist investors who are interested in the details of sectors such as biotechnology, nanotechnology or clean energy.

For Entelos, however, it was the right choice, says Karis. "Although we are in California, our customers are everywhere — Unilever and Novartis, for example. We consider ourselves an international company."

Formed in 1995 as a subsidiary of the London Stock Exchange, the AIM initially struggled to compete with the Nasdaq. But after the imposition of tough regulations on companies that list on US stock exchanges in the wake of the collapse of Texas energy company Enron, the AIM has come into its own (see Chart).

Today, 300 businesses from 26 countries other than Britain are now listed on the AIM. And although American companies make up only one-sixth of these, they are seen as a key growth area: young US companies have greater affinity for public equity markets; their European counterparts traditionally opt for bank financing.

"There is a belief now that the standard is so high to go public in the United States that it is very difficult to raise capital there if you are worth less than \$500 million," says Anne Moulier, head of US business development at the AIM. "AIM is an entrepreneurial market. It's very much designed for companies that want to grow."

Moulier concedes that companies



Open City: the rules on London's AIM stock exchange are less onerous than those on its American counterparts.

debuting on the AIM may migrate back to more established markets, such as Nasdaq, once they reach valuations of \$1 billion or \$2 billion. Still, those interviewed said, recent US entrants to the AIM are proving to be the market's best advertisements.

After Enron

The Sarbanes-Oxley corporate governance act passed after Enron's collapse in 2001 has made US public offerings more complex and expensive. On the AIM, regulatory requirements are relatively light. Companies are required to report half-yearly, as opposed to quarterly, and they don't have to provide a trading record or ensure that a minimum amount of shares are in public hands. According to financial services company Canaccord Adams, based in Vancouver, Canada, maintaining an

"The depth of understanding of biotech in Britain is not comparable with that in the United States" — Chris Hulatt

AIM listing costs about US\$900,000 per year, compared with more than \$2.3 million for Nasdaq and \$1.1 million for the Toronto Stock Exchange.

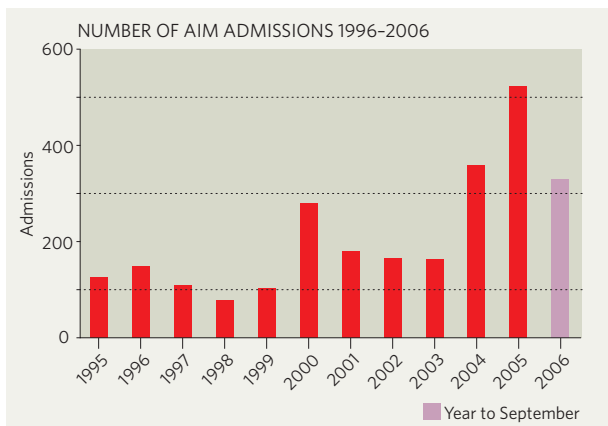
Another important difference is the unique role of a 'nominated adviser', or Nomad. This is an investment bank, or other financial institution, that shepherds a newly listed company through the flotation process, decides on its suit-

ability and determines when it is ready for an initial public offering. "The Nomad is taking a lot of responsibility," says Moulier. "It takes a lot of time and won't rush to bring a company to market." The Nomad must also force the company to release any information that could affect its

stock price — an obligation particularly appealing to large investors, she says.

Nomads help AIM-listed companies gain access to institutional investors, who make up around 95% of the investor base on the AIM, compared with some 70% in North American markets.

Opinion is divided on whether the AIM can make deeper inroads in sectors such as clean energy and biotechnology. "The depth of coverage and understanding of biotech in general in Britain is not comparable to that in the United States," says Chris Hulatt, a director of Octopus Asset Management in London. Even so, he says, "the right type of company — with a decent pipeline, strong science and management — will be able to raise money".





THE CHRISTMAS INVASION

The cheerful leaves of the poinsettia could be hiding an unwelcome visitor this festive season. **Rex Dalton** goes in search of the whitefly, a potentially devastating pest.

For many Americans, the festive season is hardly seasonal without a centrepiece of poinsettias. But for some entomologists the annual flood of red foliage is not such a welcome sight. When Timothy Dennehy lifts up a potted plant in a store or a nursery to inspect it — as he will be doing more and more over the next couple of months — he's not admiring the shape or the colour with an eye to how it might look sitting in his home. He's looking for little greenish-white harbingers of agricultural chaos.

Every autumn, Dennehy, an entomologist at the University of Arizona, Tucson, hits the state's flower stalls searching for whitefly (*Bemisia tabaci*) on the poinsettias shipped in for the festive holidays. In December 2004, his horticultural gumshoe work paid off in a Tucson market with the first US identification of the whitefly variant in question — the pesticide-resistant Q-biotype¹. The following year, the Q-biotype was found in stores across Arizona, spurring a nationwide survey that found the superfly in 22 states.

In the 1990s, the B-biotype whitefly swept through North American crops, inflicting more than a billion dollars' worth of damage on farmers in the United States and Mexico; it had hitch-hiked from Israel to Florida to California, and from there it seems likely to have been spread nationwide via the imported poinsettias. Cotton crops were devastated, melons withered on stunted vines, lettuces wilted. US farmers scrambled for scientific assistance, successfully beating the pest with a new class of insect-growth regulators — such as buprofezin

and pyriproxyfen, which were rushed through the approval process — and other pest management measures.

Now the spreading of whitefly by poinsettia is at risk of repeating itself in an even more devastating way. The Q-biotype, originally observed in Spain in 1997 (ref. 2), “is resistant to every pesticide we've tested”, says Dennehy, who co-chairs a scientific panel on the pest convened by the US Department of Agriculture (USDA)³. So far, the superfly has been found only in retail shops or nurseries. But the fear that it will one day find its way into the fields is growing. “It scares me to death; it is one pest that could completely bury us,” says Larry Antilla, an entomologist with the Arizona Cotton Research and Protection Council in Phoenix.

Festive pests

America is not the only country concerned about whitefly. Some scientists rank it as one of the world's most destructive pests to crops. Robert Gilbertson, a plant pathologist from the University of California, Davis, says damage caused by whitefly makes it “the worst [agricultural pest] problem in regions of Africa, Asia and South America”. The flies' increased resistance to pesticides and indifference to drought — many actually prefer things hot and dry — make them a grave threat to crops in parts of the developing world. Elsewhere, international trade has put tomatoes in Japan, cassava in Africa and soya beans in Australia at increased risk.

In the worst infestations, the flies can form visible clouds, coating windscreens and clogging the mouths and nostrils of field workers. Not only does the fly kill flowers, vegetables and cotton, it spreads viruses that are equally deadly in plants (see ‘At the sharp end’, overleaf). The high doses of pesticides used in attempts to control them can do more collateral damage to the insects that feed on whitefly — such as ladybirds — than to the whitefly themselves.

The best way for a country to fight the whitefly is to stop them entering in the first place. But with today's one-world agriculture, there are enormous economic and political pressures that can hamper effective inspection and regulation. In the United States, for instance, there has been a heated behind-the-scenes battle over the invasion of the Q-biotype.

The cotton industry, fearing for its fields, has fought for more aggressive control methods; the ornamental flower trade, dependent on imports that could be quarantined at borders, has opposed them. The former is a \$6-billion industry, the latter is worth \$19 billion. After Dennehy and his colleagues found the Q-biotype in Arizona — a state where the B-biotype cost cotton growers \$180 million in the first half of the 1990s — representatives of the flower industry fought against plans for a nationwide survey to check for the new variant.

Dennehy, who receives some funding from cotton growers, was displeased. As he wrote to the task force on the subject at the USDA:

“It scares me to death; it is one pest that could completely bury us.”
— Larry Antilla

M. CASTILLO/AP

"Holding off on surveys is tantamount to promoting ignorance over enlightenment regarding the biotype issue." Meanwhile, Joe Chamberlin, a North Carolina-based entomologist representing ornamental flower growers, told the USDA task-force members that his clients were concerned that the survey's "primary purpose is not to help manage the whitefly, but instead to justify imposition of regulatory action and to generate research dollars or publications for university scientists". As in the B-biotype outbreak, there has been pressure from poinsettia growers to play down their product's links to the fly. Pressure from one flower grower led to a researcher at the University of California, Riverside, having to refer to "silverleaf whitefly" rather than "poinsettia whitefly" in public statements.

On the fly

Now a nationwide survey has been established, thanks in part to the diplomatic skills of Osama El-Lissy, an entomologist at the USDA's plant-inspection service in Riverdale, Maryland. "El-Lissy masterfully cajoled many key people," says Dennehy. For his part, El-Lissy stresses that the different affected industries — cotton, ornamental flowers and vegetables — "must work together to combat this pest." As a cautionary tale, he points to the spread of the B-biotype, introduced into Florida in 1986 (ref. 4). "Back then, it was a disaster," says El-Lissy, who was conducting research in Arizona at the time. "Why? Because the industries didn't work together."

The Arizona team's investigations suggest that the Q-biotype came into the country not through Florida but through California. Poinsettia cuttings taken in central America, mostly in Guatemala, Honduras and southern Mexico, are shipped north to be grown up into plants. The Q-biotype-infected poinsettias found in Arizona were traced to a grower near Salinas, California, then south to Ecke Ranch in Encinitas, California, which handles the cuttings for a massive 70% of the United States' seasonal poinsettias.

A family-owned operation run from an old, plantation-style California home, Ecke Ranch is spread over about two dozen hectares in foothills just inland from the Pacific, largely surrounded by newly built homes for San Diego commuters. To Paul Ecke, who has run the farm in recent years, worries about the Q-biotype are "overblown".

"These guys in Arizona went off the deep end, paranoid because of the cotton growers," says Ecke. To undertake a nationwide survey, Ecke says, "was a paranoid reaction that wasn't very effective". Ecke adds that integrated pest-management controls permitted his firm to



Plant watch: Timothy Dennehy scours poinsettias in Arizona for signs of whitefly (below) in a bid to arrest the spread of this highly destructive pest and so protect US crops.



ship poinsettia cuttings last spring into the United States "with zero whitefly problems", and that the clean-up operation after the Q-biotype was first reported at his Encinitas headquarters was successful. Although his company does not publish results of its internal tests in California or Central America, Ecke says he believes things are now under control. "The Q-biotype whitefly is a non-existent problem."

Farmers must hope he is right, as there are no restrictions on Ecke — or any plant grower from California to Florida — continuing to import plants from whitefly-affected regions, even those with the Q-biotype variant. The USDA takes no action to block the import or movement of plants infected with *B. tabaci* because it has determined that the insect meets the agency's standard for a pest that can't be

regulated — which is to say, one that can't be contained or eradicated. Other nations have taken a more proactive stance. In Australia, the authorities scrupulously watched for the B-biotype after it hit the United States, although it still sneaked into the country in 1994. Earlier this year, the European Union augmented its controls on whiteflies, which have been in effect since 1992. But some Mediterranean countries still have infestations.

US flower importers fear they will lose their crops if attempts are made to regulate whiteflies because of the Q-biotype. Because the genetic analysis needed to distinguish the Q-from the B-biotype can take weeks⁵ — the two can't be distinguished visually, even under a microscope — individually packaged poinsettia cuttings being imported would die, because

N. LANEY

S. AUSMUS/USDA; A. & L. DETRICK/GRANT HEILMAN PHOTOGRAPHY

AT THE SHARP END

Like a parent with a sick child, the rancher brought the stunted tomato plant to Jesús Méndez Lozano for help. "Why are my plants dying?" he asked.

With the equipment in his small lab, Méndez, who works for the National Polytechnic Institute in Guasave, Mexico, found that there was foreign DNA in the sick plant. This DNA had the same molecular weight as that of the dreaded tomato yellow leaf curl virus (TYLCV) — indigenous to Egypt⁶ and first found in the Western Hemisphere in 1994 in the Dominican Republic. Méndez sent the DNA to a larger lab for sequencing. Two weeks later, his fears were confirmed.

Thanks to the whitefly, a pest farmers on Mexico's Pacific coast have been struggling with for 15 years, TYLCV had been provided with a route up from Central America. Now it had invaded the largest winter tomato producing region in North America.

"I will never forget that day" last November, says Méndez. Neither will the ranchers of Sinaloa, the state in which Méndez is based and where the tomato is so important that its picture appears on vehicle number-plates.

Agricultural authorities estimate



Whitefly spread the dreaded tomato yellow leaf curl virus to Mexican plantations.

that last winter about 75% of the state's tomato crop was wiped out by the virus, with the effects getting worse as the season went on. Ranchers who had started off losing 30% of the first planting were losing 100% by the third. They didn't even try a fourth.

There is no cure for the virus, and the cost of pesticides to kill the whitefly — the B-biotype, in this case — is too much for many small farmers. Last winter's economic damage was so severe, a drive through the state shows ranchers often can't afford to plough over old plants or weeds, a failing that allows the virus and its whitefly host to get better established. In 2006, as part of an integrated management plan, the summer

planting season for soya beans in Sinaloa was largely halted, thereby denying the whiteflies sustenance. In 2005, federal subsidies for soya beans — provided despite the objections from tomato ranchers — is thought to have worsened the effects of the tomato virus.

To know how to treat this sort of viral infection, a farmer

has to distinguish between different types of gemini virus — the family to which TYLCV belongs⁷. Méndez's lab at Guasave, a prosperous community on Mexico's main north-south highway, is an example of how the necessary expertise can be provided. In the five years since he and his researcher wife Norma Leyva Lopez arrived, they have upgraded the facility to offer the most advanced analysis, such as DNA amplification while-you-wait.

Realizing the importance of such knowledge, nearly 500 ranchers and businessmen came to a scientific conference on the control of whitefly and viruses that was held in late August in

the nearby city of Mazatlan.

"Ranchers will come in with a scientific article and ask: 'Can you do this?'" says Méndez, who did postdoctoral studies at the Scripps Research Institute in La Jolla, California. "I say, 'Yes. But we need more facilities'." To this end, Méndez is building support for a new agricultural biotechnology lab.

South America, Africa and Asia all face similar challenges as they seek to offer on a worldwide basis niche crops that will bring high prices when they are not available from traditional growing regions. Stations such as Méndez's are the outposts with which scientists from developed regions must work to defeat invasive pests.

One reason for such collaboration is that the pests, while limited by available food and climate, do not respect national borders. TYLCV has recently been identified in tomatoes in southern Texas, says Robert Gilbertson, a plant pathologist at the University of California, Davis. "I don't mean to be Chicken Little and say the sky is falling, but we are getting worried," he says. "There are no US controls on plant movements. It takes only one tomato transplant to get the virus started." **R.D.**

they need water and nourishment that couldn't readily be provided for quarantined products.

Under the strain

Meanwhile, the B-biotype still presents problems of its own. Lance Osborne, an entomologist at the University of Florida's research facility in Apopka and Dennehy's co-chair on the USDA science panel, points out that it is widespread in crops in the southeast, from cotton in Georgia to ornamental flowers and vegetables in Florida. It takes just two or three B-biotype whiteflies to do their damage to a plant, sucking out sap and leaving leaves with silver-coloured sooty mould; the Q-biotype seems to require hundreds of flies to do equivalent damage, Osborne says.

Osborne's worry is that farmers could promote resistance in the B-biotype by using more pesticides — as they may well do if the Q-biotype gets into fields. "I'm very concerned the B-biotype will become pesticide resistant," he says. "The B-biotype fly lays more eggs than the Q; it

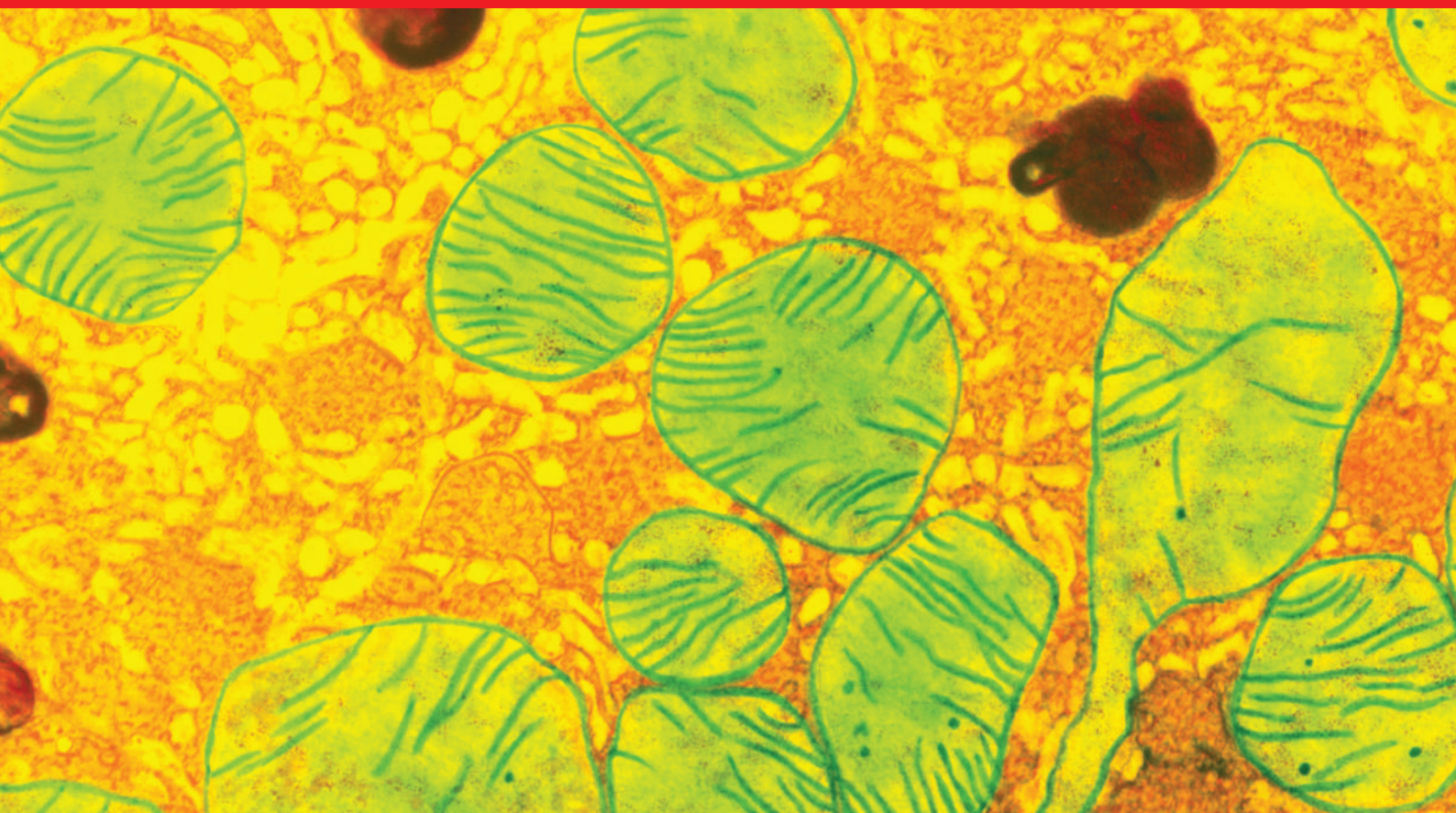
could be more of a superfly than the Q-biotype." To avoid over-reliance on pesticides, Osborne and other scientists champion integrated pest-management programmes. Along with distinct applications of the pesticides, these also involve denying the flies food by not planting some crops in hot weather, and devoting significant amounts of land to crops the flies won't infect — they love melons, peppers and leafy vegetables, but not alfalfa, wheat or corn.

But these courses can cause economic strain. In the Mexican state of Sonora, where the Yaqui River Valley has produced bumper grain crops, farmers have largely halted the planting of soya beans that mature in late summer. With fields fallow, revenue and jobs disappear. Without summer crops with labour-intensive harvests, many of the poorest Mexicans lost their jobs, transforming rural cultures as people moved to cities looking for work. In lower Sonora alone, agricultural economists estimate the losses from not planting their 70,000 hectares of summer soya beans at about \$75 million.

Now entomologists are looking at the possible responses to the pesticide-resistant Q-biotype, such as management schemes that make use of new members of the neonicotinoid family of pesticides. Far better, though, if the whitefly could indeed be kept out in the first place, however pretty its Trojan horses. ■

Rex Dalton is Nature's West Coast Correspondent

1. Dennehy, T. J. *et al.* *Resistant Pest Management Newsletter* **15** (2), http://whalonlab.msu.edu/rpmnews/vol15_no2/globe/Dennehy_et_al.htm (2006).
2. Guiraro, P., Beitia, F. & Cenis, J. L. *Bull. Entomol. Res.* **87**, 587–593 (1997).
3. Dennehy, T. J. *et al.* in *2005 Vegetable Report* (eds Byrne, D. N. & Baciewicz, P.) http://ag.arizona.edu/pubs/crops/az1382/az1382_2.pdf (Univ. Arizona, Tucson, 2006).
4. Osborne, L. S., Hoelmer, K. A. & Yokomi, R. K. in *Sweetpotato Whitefly Mediated Vegetable Disorders in Florida* (eds Yokomi, R. K., Narayanan, K. R. & Schuster D. J.) 49–52 (Inst. Food Agric. Sci., Univ. Florida, Miami, 1990).
5. Brown, J. K. *Virus Res.* **71**, 233–260 (2000).
6. Nakhla, M. K., Maxwell, D. P., Martinez, R. T., Carvalho, M. G. & Gilbertson, R. L. *Plant Dis.* **78**, 926 (1994).
7. Rojas, M. R., Hagen, C., Lucas, W. J. & Gilbertson, R. L. *Annu. Rev. Phytopathol.* **43**, 361–394 (2005).



POWER GAMES

There's a fight going on inside all our cells for each breath of air. **Nick Lane** sheds therapeutic light on the implications for cancer and degenerative diseases.

Seventy-five years ago, Otto Warburg's star was at its zenith. The pioneering German biochemist delivered his Nobel address in December 1931. He described the ingenious experiments by which he had unmasked the enzyme responsible for the critical step of cell respiration, the process that turns the energy in chemical compounds into energy the cell can use. His work on respiration in the early 1930s nearly earned him a second Nobel, ultimately denied him by Hitler. Then his star began sinking. His ideas on the importance of cell respiration in cancer led many to dismiss him as a crank. And the rise of molecular genetics in the 1960s put such ideas into a far distant orbit.

But now, Warburg's star is rising again. A new generation of researchers is returning to his ideas about respiration in cancer cells. Recent findings suggest that the enzyme he identified, cytochrome oxidase, is a key player in a new understanding of how the cell's energy metabolism affects health and disease. And surprisingly they show that light has a profound effect on how the enzyme works — and could even be used to treat degenerative disease.

To extract energy from molecules, the cell first breaks down glucose into simpler mole-

cules via a process called glycolysis. It then feeds these molecules into energy-producing structures called mitochondria, which strip electrons from them to produce energy with the help of oxygen. As Warburg showed, cytochrome oxidase governs the last reaction in this process.

Perhaps the most surprising aspect of the renaissance of Warburg's ideas is that the methods he used to make this discovery matter again. They exploit two chemical quirks: carbon monoxide (CO) can block respiration by binding to cytochrome oxidase in place of oxygen; and a flash of light can displace it, freeing up the site for oxygen to bind again.

Self control

By measuring oxygen consumption at different wavelengths of light, Warburg worked out that the enzyme belonged to a group of proteins that include haemoglobin and chlorophyll. But for Warburg, the binding of CO to the enzyme was just an oddity he could put to good use. He had no inkling that biology might use the same trick.

Yet over the past decade, researchers have come to appreciate that cells often use CO, and to an even greater extent NO (nitric oxide), to block respiration. Not only that, but light has striking counter-effects on cytochrome oxidase. And all these suitors to the enzyme turn out to be critical to our understanding not just of cancer, but practically all degenerative diseases.

Nitric oxide is emitted by nerve endings and can act on an enzyme called guanylate cyclase to relax blood vessels — the impotence drug Viagra manipulates this system. For a long time, scientists thought that guanylate cyclase was NO's only target. But in the mid-1990s, they found that the

molecule can also bind to cytochrome oxidase and hinder respiration¹. What's more, it could do so at levels found in the body's tissues.

The finding that the body could poison one of its own enzymes was initially shrugged off as an imperfection — an example of how evolution cobbles organisms together with no forethought. But a few years later, several groups reported that mitochondria harboured

"The finding that the body could poison one of its own enzymes was initially shrugged off as an imperfection."

S. GOSCHMEISSNER/SPL

an enzyme that synthesizes NO (ref. 2). Why would cells go out of their way to cook up NO right next to the respiratory enzymes?

According to cell biologist Salvador Moncada of University College London, evolution really has crafted cytochrome oxidase to bind not only oxygen but also NO. "One effect of slowing respiration in some locations is to divert oxygen elsewhere in cells and tissues," he says. This prevents oxygen levels sinking dangerously low. Moncada, for example, has shown that NO blocks respiration in the cells lining blood vessels and that this helps to transfer oxygen into smooth muscle cells in these vessels. Fireflies use a similar trick to flash light (see 'Flashing the gauntlet').

But the consequences of blocking respiration go beyond diverting oxygen. Respiration, Moncada argues, is not just about generating energy, it's about generating feedback that allows a cell to monitor and respond to its environment. Blocking respiration generates chemical signals, in the form of highly reactive molecules called free radicals. These are normally associated with cell damage, but now it seems they can interact with the proteins that control gene activity and adapt cells to changing circumstances. "Free radicals had a bad reputation," says pathologist Victor Darley-Usmar, of the University of Alabama, Birmingham. "But we now see them as signals."

In the past few years, researchers have compiled a list of these proteins, or transcription factors, the activity of which depends, at least in part, on interactions with free radicals³. These

include many proteins known to be linked to cellular life and death, such as P53, a protein that kills cells if they show signs of turning cancerous. "The question is," says Darley-Usmar, "how is the whole system controlled?"

The answer lies in the tens of thousands of protein machines that line each mitochondrion. Organized into chains, they pass electrons extracted from broken-down glucose on to each other. The end-point of the chain is cytochrome oxidase, which catalyses the final step of respiration, in which electrons and protons are transferred onto oxygen to form water. Energy released by this process is used to pump protons over a membrane, creating an electrical charge. The cell can then use this charge to power an enzyme that makes ATP, a molecule that fuels chemical reactions in the cell.

The cell can suppress the number of free radicals coming from these respiratory chains by allowing protons to leak back through the membrane without driving the synthesis of ATP, a process known as uncoupling⁴.

Undercover radicals

But if uncoupling doesn't bring free-radical leak under control, the signal may be amplified. "There's a cross-talk, known as the retrograde response, between the mitochondria and genes in the nucleus, which we're only just beginning to explore," says cell biologist Keshav Singh of the Roswell Park Cancer Institute in Buffalo, New York, who helped uncover this mechanism. "If we can modulate this signalling, we might be able to influence the life or death of

"If we can modulate this signalling, we might be able to influence the life or death of cells in many pathologies."

— Keshav Singh

Gas station: by binding to the cytochrome oxidase enzyme (left), NO regulates respiration.

cells in many pathologies, even ageing."

In cancer cells, something goes awry with this pathway. Free radicals can trigger two steps in carcinogenesis, chromosomal instability and metastasis⁵, by hindering DNA repair and promoting the activity of genes involved in cancer spread. "We've found that damaged mitochondria, which overproduce free radicals, under-

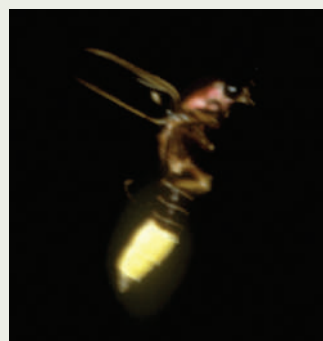
Flashing the gauntlet

Firefly flashes have lit up the human imagination since antiquity. Light is produced by the action of the luciferase enzyme, which uses ATP to activate luciferin to a luciferyl adenylate intermediate. This in turn reacts spontaneously with oxygen to emit a photon, giving rise to bioluminescence.

The sequence of flashes, which often beam complex messages to potential mates, clearly requires exquisite nervous control. But no nerve cells connect directly to the 'lantern' cells that produce the light, suggesting that some kind of intermediary is needed to ferry messages from the nerve endings.

Barry Trimmer and his colleagues at Tufts University in Medford, Massachusetts, have shown that this is likely to be NO (ref. 12) and suggest it works by binding to cytochrome oxidase¹³. To do so, it exploits a neat feature of the firefly's breathing system.

Insects breathe through tubules, called trachea, that connect the outside world to cells in the body via a network of increasingly fine branches. The ends of these, called tracheoles, abut cells and allow oxygen to diffuse outwards. In fireflies, when oxygen passes from tracheoles into the lantern cells, it must run the 'gauntlet' — a



dense ring of energy-producing structures called mitochondria in the periphery of the lantern cells.

Normally, the oxygen would bind to cytochrome oxidase in

the mitochondria and be used in respiration. But when NO is released from the nerve endings, it binds to cytochrome oxidase, blocking respiration and diverting oxygen through the gauntlet into the lantern cell.

There, oxygen reacts with the luciferyl intermediate to produce a flash of light. The glory is that the flash switches itself off. Light dissociates NO from cytochrome oxidase, allowing oxygen to bind again. Now the mitochondria consume oxygen once more, allowing the luciferyl intermediate to build up until another wave of NO arrives.

N.L.

F. LEROY/BIOCOSMOS/SPL

J. E. LLOYD/PHOTOLIBRARY.COM



mine the integrity of the nuclear genome in otherwise normal cells," says Singh.

Ultimately, however, the most intractable problem with cancer cells is their failure to undergo apoptosis (cell suicide). There are echoes of Warburg here, for he argued that cancer cells revert to a primitive type that has little or no need for oxygen or mitochondria, depending instead on anaerobic forms of energy generation such as glycolysis. Whether cells really 'revert' is a moot point, but Moncada has shown that cells' ability to switch to glycolysis is critical to their survival⁶.

Moncada's findings suggest that cells that can do without mitochondria — many stem cells, for example, which have been implicated in cancer — can have their resistance to apoptosis stiffened by NO binding to cytochrome oxidase, making cancer more likely. At the same time, cells that depend on mitochondria for energy, such as neurones, may be pushed to apoptosis by NO binding, making degenerative disease more likely.

Any solution to excessive NO binding might lower the risk of both cancer and degenerative diseases, as it would make apoptosis more likely in cancer cells and less likely in normal cells. And the second essential feature of Warburg's experiments — light — might do just that.

Light has long been known to promote wound healing, but the detailed molecular mechanisms have only recently been studied. Light's effects are more than skin deep: at

long wavelengths, in the near infrared (NIR) spectrum, photons may penetrate several centimetres into the body.

And according to photobiologist Tiina Karu, at the Russian Academy of Sciences in Moscow, NIR rays in exactly this range modulate cell respiration and the signals it generates.

Light works

Karu's group has explored a number of changes taking place in cultured cells in response to NIR rays. The immediate effect is an energy 'buzz', in which ATP levels are stoked up and the electrical charge across the mitochondrial membrane is strengthened. A few hours later, the activity of as many as 110 genes shifts in concert⁷. These genes orchestrate a prolonged rise in mitochondrial energy production, as well as stress resistance. They also prompt cells to cling more strongly to their surroundings, an important factor in wound healing⁸.

Exactly how NIR light interacts with the enzyme to bring about these changes is unclear, and difficult to measure.

One idea is that phototherapy might work by dissociating NO from the enzyme, so reversing the signalling consequences of excessive NO binding.

"We have shown that light can indeed reverse the inhibition caused by NO binding to cytochrome oxidase, both in isolated mitochondria and in whole cells," says biochemist Guy Brown, at the University of Cambridge, UK. "And what's more, we found that light can protect cells against NO-induced cell death."

He has a reservation, however. These experiments used light in the visible spectrum, with wavelengths from 600 to 630 nm. Although Brown acknowledges that cytochrome oxidase absorbs NIR photons from 700 to 900 nm, he points out that the absorption takes place in part of the enzyme not involved in NO binding. NIR also seems to have effects on cytochrome oxidase in conditions where NO is unlikely to be present. Presumably then, says Brown, NIR must also have a direct effect on the enzyme.

Regardless of the exact mechanism, the effect is to relieve a blockade of the enzyme, whether by NO or any other molecules. According to toxicologist Janis Eells, at the University of Wisconsin, Milwaukee, such relief lowers the likelihood of apoptosis in many conditions.

Eells and her colleagues found that NIR phototherapy counters methanol poisoning,

which injures the retina and optic nerve, often causing blindness⁹. The toxic metabolite is formic acid, which inhibits cytochrome oxidase. "In a rat model, NIR phototherapy is able to restore virtually normal retinal function, at least as judged by the electroretinogram," says Eells.

And neurobiologist Margaret Wong-Riley and her colleagues at the Medical College of Wisconsin in Milwaukee have shown that NIR phototherapy can also oppose the effects of cyanide on cell cultures¹⁰. Cyanide poisons by binding to cytochrome oxidase. Wong-Riley's team showed that phototherapy could halve the rate of apoptosis in cultured neurones, even when given before cyanide treatment.

But can NIR phototherapy relieve not just acute toxicity, but more chronic inflammatory conditions? The signs augur well. Eells and her colleagues have shown that NIR phototherapy could cut the rate of apoptosis by 50% in a rat model of retinitis pigmentosa, in which photoreceptors die by apoptosis during postnatal development causing retinal degeneration and blindness¹¹.

The use of NIR phototherapy on other conditions in which cells die by apoptosis, including acute ischaemic stroke and myocardial infarction, has also shown promise in animal models, and should soon enter clinical trials.

So will it work? "The trouble is that if you enthuse about light as a therapy, clinicians and even researchers tend to back away," says Eells. "Perhaps if the physical interactions of photons with

cytochrome oxidase and NO are better known, people will begin to appreciate the huge potential benefits of this technique."

Nick Lane is a science writer based in London.



Healing glow? Light can prevent cells from killing themselves.

1. Moncada, S. & Erusalimsky, J. D. *Nature Rev. Mol. Cell Biol.* **3**, 214–220 (2002).
2. Brown, G. C. *Biochim. Biophys. Acta* **1504**, 46–57 (2001).
3. Darley-Usmar, V. *Free Rad. Biol. Med.* **37**, 753–754 (2004).
4. Brand, M. D. *Biochem. Soc. Trans.* **33**, 897–904 (2005).
5. Singh, K. K. et al. *Gene* **354**, 140–146 (2005).
6. Moncada, S. & Bolaños, J. P. *J. Neurochem.* **97**, 1676–1689 (2006).
7. Zhang, Y. et al. *J. Invest. Dermatol.* **120**, 849–857 (2003).
8. Karu, T. I., Pyatibrat, L. V. & Afanasyeva, N. I. *Photochem. Photobiol.* **80**, 366–372 (2004).
9. Eells, J. T. et al. *Mitochondrion* **4**, 559–567 (2004).
10. Wong-Riley, M. T. et al. *J. Biol. Chem.* **280**, 4761–4771 (2005).
11. Eells, J. T. et al. *Near Infrared Light Therapy for Retinitis Pigmentosa* (ARVO, 2006).
12. Trimmer, B. A. et al. *Science* **292**, 2486–2488 (2001).
13. Aprille, J. R., Lagace, C. J., Modica-Napolitano, J. & Trimmer, B. A. *Integr. Comp. Biol.* **44**, 213–219 (2004).

Drugs from the deep

Is the cure for cancer lurking beneath the waves? **Emma Marris** plunges into the chemistry of marine natural products.

Walk downhill from the University of California, San Diego, towards the ocean, and you'll see bungalows draped in bougainvillea, tanned youths unloading surfboards from station wagons, and at the bottom, William Fenical's lab at the Scripps Institute of Oceanography. The chemist has white sand and palm trees both on his screensaver and outside his window. And beyond them — the sea.

Fenical and a global group of like-minded scientists believe that the sea hides a mermaid's grotto of useful chemicals. And they have been diving for corals, grinding up sponges, fermenting microbes, and poking around inside cells for decades in an effort to find them. Chemicals from these organisms — natural products — can have pharmaceutical potential. But before scientists could deliver any useful natural products, the drug industry largely lost interest in the field, slowing its growth. Fenical and others now say that the problems of the past have mostly been solved. And they may no longer need the drug industry's support.

Many of the best-known drugs are extracted from living things. A famous example is the actinomycetes, soil microbes so good at producing antibiotics that in the 1950s and beyond, pharmaceutical companies sent people tramping all over the globe to collect dirt. Most of our antibiotics still come from *Penicillium* mould or soil microbes. "Antibiotics were an astounding discovery," says Fenical from his seaside office, "without question the most important medical discovery in history."

The ocean covers two-thirds of Earth, and many branches of life — such as the group containing jellyfish and corals — are exclusively marine. Yet until relatively recently, chemists were confirmed landlubbers. "We are not marine organisms," says Fenical, "so until about 1970, no one even thought of the ocean. It was left as a deep secret." Fenical, and a few others, were the first to get their toes wet. "It seemed ridiculous to me that the ocean — with such a vast habitat — had escaped anyone's notice," he says. "But there are good reasons. People fear the ocean; it has been considered a very hostile, inhospitable place."

Thirty years later the field is having a mini renaissance¹, with strong research labs around the globe, including China. But natural-products chemists have struggled to convince

pharmaceutical companies to share their enthusiasm. Most drug companies dismantled their natural-products research units in the 1990s, when combinatorial chemistry was heralded as the next wave of drug discovery. At that time, there were no drugs based on marine natural products on the shelves. Today, there is one: Prialta, a pain reliever derived from the hunting venom of the cone snail (see 'Marine medicines'). "Not to say that there aren't powerful leads in nature," says Kate Robins, a spokeswoman for Pfizer in New London, Connecticut, which closed its natural-products programme in the 1990s, "but it has been our experience that promising leads come faster and more frequently out of combinatorial chemistry and synthetic techniques."

Backwater

Fenical bemoans what he perceives as the drug companies' lack of vision. "We have done so much to prove the rich resource of the ocean. The big pharmas are apathetic and uninterested. They are risk averse." This irks him, he says, not just for professional reasons. "The fact is that no really new antibiotic has come out since the late 1980s. We are on a very dangerous collision course with a plague."

Marine natural products are attractive sources for new antibiotics because they are mostly secondary metabolites. That is, they are not essential to an organism's growth and development, but are compounds that do something else, such as deter predators — and could be re-engineered to aid our fight against infectious disease.

But infectious diseases are unfashionable with many drug companies, and the legal uncertainties around bioprospecting for natural products hasn't helped its cause. Another problem is scarcity. A compound made in tiny



"We have done much to prove the rich resource of the ocean. The drug companies are apathetic and risk-averse."
— William Fenical



quantities by a sponge that lives hundreds of metres down, for example, poses challenges to traditional drug testing and development.

So scientists are now working out how organisms make natural products. To their surprise, many compounds are generated by an associated microbe, rather than the organism itself. This means they can either grow the microbes alone in flasks, or they can identify the microbial genes responsible for making the compound, and insert them into an organism researchers find easier to work with, such as *Escherichia coli*. In addition, more and more natural products are being made from scratch as synthetic chemists get better at their art².

David Newman, director of the natural-products division at the US National Cancer Institute in Bethesda, Maryland, sums it up: "In the 1980s, the field was hyped, but we had not solved the supply problem. By the end of the 1990s, what came together was molecular biology, which allowed you to go looking for the producing genes, the ability to grow microbes, and synthetic chemists who could either do total synthesis or work from a structure."

Fenical's research followed a similar pattern. He previously studied invertebrates on coral reefs, where low nutrient levels create competition for food and may favour the evolution of toxins. Resilience, a skin cream made by Estée Lauder (selling for \$50 an ounce), contains an

SCRIPPS



PHARMAMAR

Treasure hunters: divers from biotech company PharmaMar collect marine life in the search for drugs.

anti-inflammatory that Fenical found in soft corals during those years. But he eventually left the invertebrate field: "It was almost impossible to develop a drug, because we could not provide lots and lots of it."

Instead, he turned to microorganisms. With patience and luck, marine microbes are fermentable — they can be grown in large flasks, if the right mix of nutrients is provided.

The lab made its first big discovery in deep-ocean sediment. There the team found lots of actinomycete species, despite the received wisdom that there were none in the sea. The team named the first genus of these *Salinospora*³. And from one of them, in 2003, came salinosporamide A, a compound that binds extremely selectively to the proteasome in tumour cells⁴. It is now in clinical trials for multiple myeloma, a cancer of the blood.

Recently, the lab has discovered another new genus, *Marinispora*, which produces compounds with promising antibiotic and anticancer properties⁵. Fenical says the limiting factor now is time and the logistics of getting samples from deep-sea mud.

Working with Scripps technicians, Fenical has devised probes that fall to the sea floor, drive a corer into the mud, release weights and rise to the surface. The samplers take 45

minutes to fall 6,000–7,000 metres to the ocean floor and return — if they come back at all.

Technical advances are changing the field in other ways. Studies using nuclear magnetic resonance imaging need only tiny quantities of compounds, making the work easier and more ecologically sound, says John Blunt, a chemist at the University of Canterbury in Christchurch, New Zealand. "The other area that is developing is assays," says Blunt. "We used to use simple cytotoxic and antibiotic assays. Now specific enzymes are being used." So instead of finding out whether a compound has any biological activity, researchers can find out whether it binds to a particular target.

Find a niche

Whether big drug firms renew their interest in marine products may not matter — small institutional labs or companies can take them on. "Many of our big-pharma competitors who have moved out of this effort are getting back in through collaborations with smaller companies," says Guy Carter, vice-president of research at drug firm Wyeth in Pearl River, New York, which, unlike most drug companies, retains a natural-products unit employing 50 people.

For 20 years, the Madrid-based biotech

Marine medicines

Ecteinascidia turbinata, a

Caribbean and Mediterranean sea squirt, makes a compound that PharmaMar has brand-named Yondelis. The firm is awaiting marketing approval from the European Medicines Agency for this antitumour compound for soft-tissue sarcomas. Trials are also under way for ovarian cancer.



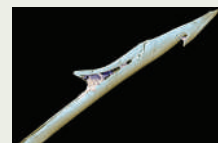
Aplidium albicans. Another

anticancer drug has been isolated from this sea squirt. PharmaMar calls it Aplidin and is testing its efficacy.



Conus magus.

This cone snail paralyzes its prey using a poison-tipped barb (right). The poison is a painkiller many times more potent than morphine, and is now on the market as Prialt.



PHARMAMAR

PHARMAMAR

V. STEGER/SPL

company PharmaMar has sought cancer drugs in marine natural products. It investigates compounds found by academics, and has its own explorers. It now has a library of more than 40,000 compounds, six of which are in clinical trials. Fenical helped found a spin-off, called Nereus, after the Greek god of the Mediterranean. The eight-year-old company has two compounds in preclinical trials for cancer.

Despite the scarcity of new drugs on the market, marine drug discovery still attracts newcomers. Robert Jacobs, a marine natural-products biologist at the University of California, Santa Barbara, has a tip for new recruits. "When we started out in the early 1970s, just about anything you picked up, you'd find something new. These days we recommend that people think ecologically." That is — find a niche where secondary metabolites are likely to be abundant and focus on it. In the vastness of the oceans, there's still much to explore. ■

Emma Marris is a reporter for Nature based in Washington DC.

1. Blunt, J. W. et al. *Natural Prod. Rep.* **23**, 26–78 (2006).
2. Nicolaou, K. C. & Snyder, S. A. *Proc. Natl Acad. Sci. USA* **101**, 11929–11936 (2004).
3. Mincer, T. J., Jensen, P. R., Kauffman, C. A. & Fenical, W. *Appl. Environ. Microbiol.* **68**, 5005–5011 (2002).
4. Feling, R. H. et al. *Angew. Chem. Int. Edn* **42**, 355–357 (2003).
5. Kwon, H. C. et al. *J. Am. Chem. Soc.* **128**, 1622–1632 (2006).

RNAi Nobel ignores vital groundwork on plants

SIR — The Nobel prize, by recognizing the individuals behind breakthroughs, inspires all scientists to do great science. The discovery of RNA interference (RNAi) changed the face of gene regulation, a feat deservedly recognized with this year's Nobel Prize in Physiology or Medicine¹.

As undergraduates, we witnessed with great excitement the discovery of gene silencing. At that time, almost all research in that area was being conducted by plant scientists, and as young plant biologists we were lucky to have front-row seats to this molecular drama.

Like all great advances, RNAi is turning out to be important in ways that could not have been guessed at even a decade ago. Therefore we were not surprised to discover that the topic was selected for this year's honour — but we were shocked that the plant scientists who were so crucial in discovering and communicating the underlying mechanism of RNAi were not awarded a share.

Of course there is often controversy around the awarding of the Nobel prize. Yet in this case we feel that a grave error has been made in overlooking key researchers, all of whom work on plants. Most of the six points cited in support of the prize were not first shown by Andrew Fire or Craig Mello, who won the prize, but were already known from plant research. For example, the sequence specificity, RNA degradation and post-transcriptional nature of gene silencing had all been shown in studies on plants and plant viruses^{2,3}. In addition, the observation that silencing is non-cell-autonomous was first done in plants⁴. Moreover, the models involving double-stranded RNA and amplification mechanisms had been proposed by plant researchers before the publications of RNAi mechanisms in animal systems⁵.

In our view, the main importance of the work by Fire, Mello and colleagues (accessible via ref. 1., together with other relevant articles) was the integration of these elements to demonstrate that they stood up to testing in an animal system, the nematode worm *Caenorhabditis elegans*. Subsequently, plant research continued to break new ground on mechanisms of RNAi-based genetic regulation.

As the Nobel prize may be shared by three people, a plant scientist should have been included. One who springs to mind as a pioneer in the field is David Baulcombe (see www.sainsbury-laboratory.ac.uk/dcb). His work was key to understanding the mechanism of RNAi and paved the way for Fire and Mello's findings.

By ignoring the work done in plants, the Nobel committee has undermined the

values at the centre of the prize and is sending a discouraging message, especially to young researchers.

Marc Bots*, **Spencer Maughan†**,
Jeroen Nieuwland†

*Flanders Interuniversity Institute for Biotechnology, Technologiepark 927, BE-9052 Gent, Belgium

†Institute of Biotechnology, University of Cambridge, Cambridge CB2 1QT, UK

1. *Nature* **443**, 488 (2006).
2. Baulcombe, D. C. *Plant Mol. Biol.* **32**, 79–88 (1996).
3. Van der Krol, A. R. et al. *Plant Cell* **2**, 291–299 (1990).
4. Voinnet, O. & Baulcombe, D. C. *Nature* **389**, 553 (1997).
5. Metzlaiff, M., O'Dell, M., Cluster, P. D. & Flavell, R. B. *Cell* **88**, 845–854 (1997).

Iran seeks nuclear power to replace reliance on oil

SIR — I read with interest your Editorial “Revival in Iran” about the development of the country's education and science (*Nature* **442**, 719–720; 2006). As you say, Iran made many achievements during the time of Nizam al-Mulk. Its history also includes many other periods in which highly qualified scientists such as Razi in the ninth century and Avicenna in the tenth made great advances.

You suggest that Iran's desire for nuclear technology may have served as an impetus for its rise in science, and that this programme is viewed with suspicion by many. It is not clear to me why you doubt that growth in science will be maintained. The acquisition and use of technology, including nuclear, is a major item of Iranian scientists' agenda, in the hope that nuclear power can replace oil as the country's main source of energy.

Your Editorial states that more than 4,000 papers from Iranian scientists were published in 2005. In fact, ISI indicates that 5,576 papers were published in 2005, compared with 4,343 so far in 2006: a figure likely to rise to about 6,500 by the end of the year — a 17% increase. If research into medical and related sciences is considered separately, Medline indicates that 1,466 Iranian papers were published in 2005, whereas for 2006 to date the figure is 1,440 and likely to rise to about 2,100 by the end of the year: an increase of 46%. The trend of scientific progress is certain to be maintained and the “worryingly inexperienced people” appointed to lead the universities will steer Iran to new horizons. The average qualification of these new incumbents is higher than their predecessors — the latter were well qualified, but even better-qualified people are now available. The outgoing presidents have become members of national decision-making bodies, for example the health policy council, and sit on universities' boards of trustees. Rather than their experience being lost, it is used to help exchange views with the new appointees.

Iran has not faltered in maintaining its momentum for contributing to science, and

will not do so in the future. We believe that Iranian scientists can and will respond appropriately to the country's needs.

Kamran B. Lankarani

Ministry of Health and Medical Education of I. R. Iran, Tehran, I. R. Iran

Iran: productivity is not simple to evaluate

SIR — Eran Meshorer, in Correspondence (“Iran is sixth, not second, in Middle East publication list” *Nature* **443**, 271; 2006), states: “When the number of publications (see www.ncbi.nlm.nih.gov/entrez) is corrected for population size, Iran becomes only the sixth in terms of scientific productivity.” However, the scientific productivity of a country is not the number of publications divided by the number of people. To evaluate productivity, one has to account for other important factors, such as publication citations. In addition, the NCBI website is mainly devoted to life-sciences journals and does not cover the main journals in other fields.

We can usually use our common sense to evaluate a country's scientific productivity from what we can observe of its output: an article at http://en.wikipedia.org/wiki/Science_in_Iran gives some background. But if a more systematic approach is desired, it is certain that there is no single formula that can provide a definitive ranking.

Hamid Reza Maei

Department of Physiology, University of Toronto, Toronto, Ontario M5S 1A8, Canada

Iran: let's keep politics in the realm of rationality

SIR — I read with interest the debate about Iran's scientific development in your Editorial and Correspondence (*Nature* **442**, 719–720 and **443**, 271; 2006). The Correspondence authors argued that Iran poses a danger. However, I believe that, as scientists, we have a duty to remain rational, believing in facts and distrusting rhetoric — whether from Iran, the United States or elsewhere.

I do not believe that evidence has been produced to show Iran's nuclear-technology programme is being used for military purposes. By relying on words, and hence on belief rather than evidence, we would elevate diplomacy and politics to the level of faith, whereas history teaches us that they should be kept in the realm of rationality. Francisco Goya told us in a famous etching that “the sleep of reason produces monsters”. Let's not make that picture into a prophecy.

Ludovico Cademartiri

Department of Chemistry, University of Toronto, Toronto, Ontario M5S 3H6, Canada

COMMENTARY



PREPARING FOR THE WORST

An international data bank of nuclear explosives is needed to determine the source of nuclear materials following an explosion, argue

Michael May, Jay Davis and Raymond Jeanloz.

How soon after an attack could the source of a nuclear explosive be determined?

The ultimate terrorist attack is a nuclear explosion in a city. The likelihood of such an event is uncertain, but the consequences would be enormous. As many as a hundred thousand people could be killed, and many more left wounded or sick; several square miles of the city would be destroyed or contaminated, and the political and social consequences would be far-reaching.

Once it had happened, there would be an urgent need to determine where the nuclear explosive came from and who was responsible, and not just for retribution: a pressing concern would be to assess the chances of another nuclear detonation. We propose an international data bank of known nuclear explosive materials to aid in that process.

We also argue that once it becomes possible to trace nuclear explosives, this could in itself deter the criminal transfer of such materials from governments to terrorists. Although sub-national groups may be able to assemble a nuclear device surreptitiously, they do not have the facilities to make weapons-grade nuclear materials¹. Terrorists must therefore buy or steal what they need.

Most nuclear explosive materials are in the former Soviet Union and the United States, along with far smaller amounts in other countries, including Pakistan and North Korea. It is feared that significant quantities of nuclear weapons materials may be kept under poor security². Because the manufacture and storage of such materials are under state control, the knowledge that significant capabilities exist for attribution should encourage better security and accounting.

There already exist several small databases

similar to what we propose. But some of these are held privately, for example at the Institute for Transuranium Elements (ITU) in Karlsruhe, Germany, or by individual governments. Even the international database held by the International Atomic Energy Agency (IAEA) in Vienna, Austria, is incomplete and not designed for the event-driven rapid forensics that would be required in response to a terrorist detonation. They therefore provide limited deterrence³.

An international data bank of the type we are suggesting would include a database (containing classified and non-classified information), a physical library of nuclear explosive materials, and vetted procedures for handling both

"Exclusion of certain sources can be as important as full attribution during the period of political uncertainty after an attack."

data and samples. Cooperation of governments will be central to its success, and once the data bank is established, refusing to provide samples would focus suspicion.

How quickly can nuclear material be identified in practice? Even after a nuclear explosion most of the nuclear material — plutonium or highly enriched uranium — remains intact, as only a fraction gets converted into energy and fission products. Much of the other material in the weapon is also left behind, usually in a highly irradiated state. Although explosion debris is spread over a wide area in tiny

fragments, its radioactivity makes it detectable, identifiable and collectable, and a wealth of information can be obtained from its analysis: the efficiency of the explosion, the materials used and, most important for our purpose, some indication of where the nuclear material came from (see Table).

Detective work

What clues would such analysis rely on? Uranium isotopes, for example, vary in composition and impurities according to where the uranium was mined, and how it was processed⁴. Weapons-grade plutonium can be exposed during its production to different neutron fluxes and energies, depending on the particular reactor used. It is also possible to detect the length of time plutonium has spent in the reactor. Such information may not point to a single origin, let alone determine the full chain of transfer from source to terrorist group. Still, exclusion of certain sources can be as important as full attribution during the period of political uncertainty after an attack, and is usually more feasible.

An adequate data bank would consist of three components: a database that lists key chemical (elemental and isotopic) and physical properties of known plutonium and highly enriched uranium samples; a suite of samples or access to representative samples of these nuclear explosive materials; and validated procedures for performing the analyses and for handling the samples and information in the database. The database would collect best-estimate values and uncertainties for all measurements.

The database would need to avoid any

suspicion of political bias by conducting chemical and physical analyses in several laboratories in different countries under the international auspices of a body such as the IAEA. Such transparency is crucial both when the database is set up and during forensics activity after an attack.

A key technical issue is establishing validated methods for performing analyses or for handling the samples, and is being partially addressed by the Nuclear Smuggling International Technical Working Group at the ITU. This group of international experts is developing mutually agreed-upon techniques for performing reliable nuclear forensics, which are then peer-reviewed and regularly benchmarked in tests using unfamiliar samples⁴.

Access to physical samples is needed to validate the database, and to allow more detailed analyses in the event of a nuclear detonation. The best approach would be to archive gram or subgram quantities of representative material under the control of an international body such as the IAEA. But this may not always be possible because of concerns over sensitive military or commercial secrets — concerns that are codified in law in some states.

Secrets that could be revealed by the sort of database described here include reactor-fuel production technologies that may have economic value in the eyes of the fuel producer. Consequently, the database would probably need to have both public and classified parts. Because this could threaten confidence in the database, we advocate establishing as much transparency as possible ahead of time, in procedures as well as data.

For the public part of the database, it should be technically possible to include key information — such as the isotopic composition of source materials — that does not reveal details of how the weapons are made. Questions may arise regarding the reliability of the database, including whether fake samples may be submitted (spoofing). We believe that fakes would be open to discovery through subsequent analyses. But such questions must be addressed and resolved, at least among partners in the database, for an attribution decision to be generally credible and accepted.

Nuclear forensic activities following a terrorist explosion

Action	Timescale	Methods and limitations
Determination that detonation is nuclear	Less than an hour	Determined from yield (seismic magnitude), optical signature and presence of excess radiation above normal background (due to neutron activation and presence of fallout). Rapid classification as unambiguously nuclear can be more challenging for low yields (sub-kiloton)
Identification of fuel type and sophistication of device, and initial assessment of isotopic signatures	Hours to weeks	Limited by time needed to collect sample, bring to laboratory and prepare for adequate analysis
Complete characterization of chemical and physical signature	1 to 2 weeks	Limited by decay rates of isotopes and by need for multiple high-resolution analyses
Attribution and assessment of further threat	Hours to years	Limited by availability of relevant data bank

There is precedent for dealing with such issues in treaty-verification regimes, which often involve a mix of public and classified methods. One approach we advocate is to allow some of the database to remain classified, but in such a way that it can be interrogated at a time of need by, for example, a group of pre-approved analysts selected from participating countries.

Secret codes

In cases when sample materials are deemed too sensitive to be provided to an international archive, procedures similar to the 'challenge inspections' agreed on by signatories to arms-control agreements could be used. Challenge inspections of weapons facilities can be requested by the international inspectorate for the Chemical Weapons Convention, for example. The right to challenge inspections was embodied in the most recent IAEA protocol for inspecting nuclear-energy sites (the so-called Additional Protocol) but has not been agreed to by all parties to the Nuclear Non-Proliferation Treaty. However, such inspections could be used to make samples available under appropriate and pre-approved conditions — notably, in response to a terrorist nuclear detonation.

For the classified part of the database, there would be information protocols that allow appropriate disclosure of sensitive data (perhaps to a small set of vetted analysts) in times of need. We recommend using 'message digests' (or 'hashing') to encapsulate analytical information and to interrogate the database¹. Hashing is often used in electronic commerce to keep information secure. The advantage of hashing is that only a small fraction of the data file is encrypted and transmitted, so it is less vulnerable to inappropriately revealing sensitive information than encrypting alone. And provided that information in the secure database is stored in a predetermined format, the task of comparing it to new forensic information is straightforward.

A primary motivation for disclosure of information to this data bank, and for participation in the associated forensic analysis, would be assurance of a seat at the table

in attribution decisions and the consequent political or military steps. Although cooperation in a data bank is not spelled out as a requirement in the Nuclear Non-Proliferation Treaty, cooperation could be mandated for states receiving nuclear-related exports from members of the Nuclear Suppliers Group, as occurred when the Additional Protocol was introduced. Alternatively, cooperation could be mandated by a United Nations Security Council resolution.

A key feature of the nuclear forensics timeline (see Table) is that analytical information can be available both before a detonation — if smuggled nuclear materials are intercepted, for example — and within hours after a terrorist nuclear detonation. But attribution will require an appropriate database for interpreting the forensics information. In the current situation, obtaining this information could require months or longer after a detonation, yet there would be great pressure for rapid, actionable information, including ruling out potential sources. Establishing the data bank proposed here would greatly reduce the time between this most terrible of events and the ability to respond to it.

Michael May is director emeritus of the Lawrence Livermore National Laboratory and is at Stanford University; Jay Davis is former director of the Defense Threat Reduction Agency; Raymond Jeanloz is chair of the National Academy of Sciences' Committee on International Security and Arms Control, and is at the University of California, Berkeley.

1. National Academy of Sciences *CISAC Monitoring Nuclear Weapons and Nuclear-Explosive Materials* (National Academies Press, Washington DC, 2005). www.nap.edu/catalog/11265.html.
2. Bunn, M. & Wier, A. *Securing the Bomb 2005: The New Global Imperatives* (commissioned by The Nuclear Threat Initiative; 2005). http://bcsia.ksg.harvard.edu/BCSIA_content/documents/thebomb2005.pdf.
3. Kristo, M. J., Smith, D. K., Niemeyer, S. & Dudder, G. B. *Model Action Plan for Nuclear Forensics and Nuclear Attribution*, UCRL-TR-202675 (Lawrence Livermore National Laboratory, 2004). www.llnl.gov/tid/lof/documents/pdf/305453.pdf.
4. Moody, K. J., Hutcheon, I. D. & Grant, P. M. *Nuclear Forensic Analysis* (Taylor and Francis Group, New York, 2005).

Acknowledgement: The authors thank Ian D. Hutcheon of the Lawrence Livermore National Laboratory and his colleagues for helpful discussions.



Stopping illegal transfers of uranium or plutonium requires government cooperation.

S. R. WINTERB/GETTY

AUTUMN BOOKS

The Chomsky of morality?

A view of morality as the product of an innate mental faculty — rather like language.

Moral Minds: How Nature Designed Our Universal Sense of Right and Wrong

by Marc D. Hauser

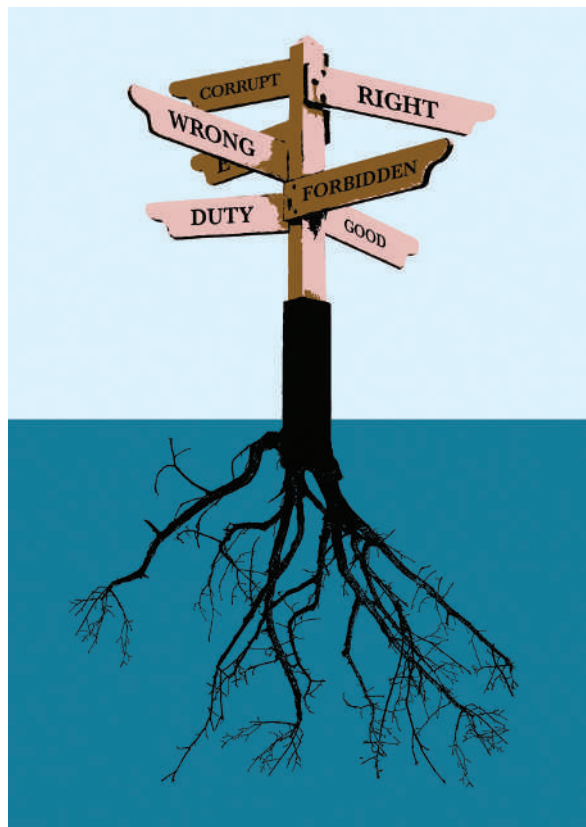
Ecco: 2006. 512 pp. \$27.95

Paul Bloom & Izzat Jarudi

In *Moral Minds*, Marc Hauser makes an audacious claim about moral thought. He argues that morality is best understood in much the same way as Noam Chomsky described language: as the product of an innate and universal mental faculty. For Hauser, moral intuition is not the product of culture and education, nor is it the result of rational and deliberative thought, nor does it reduce to the workings of the emotions. Instead, it is human nature to unconsciously and automatically evaluate the moral status of human actions: to judge them as right or wrong, allowed or forbidden, optional or obligatory.

As Hauser is careful to point out, he is not the first to make the leap from a chomskyan theory of language to a chomskyan theory of morality: this analogy was proposed by the political philosopher John Rawls, the legal scholar John Mikhail of Georgetown University in Washington DC, and by Chomsky himself. But *Moral Minds* is the first detailed exploration of this idea. It is a trade book, highly accessible to a general audience and drawing on diverse examples from literature, popular culture and history. But it is also a deeply significant intellectual contribution: everything that's done in the new science of moral psychology in the coming years is going to be a response to this important and enjoyable work.

Certain deep parallels between language and morality make Hauser's proposal worth taking seriously. Chomsky has long observed that language is a system of knowledge, but what we know (competence) is different from how we use this knowledge in everyday life (performance). Linguistic competence is also unconscious: every English speaker knows that something is wrong with the sentence "John seems sleeping," but only experts understand why. Similarly, moral intuitions are imperfectly linked to action — you can know the right thing to do but choose not to do it — and only experts can articulate adequate reasons for common-sense moral judgements. Finally, just as there are, arguably, innate principles of language, Hauser reviews an extensive array of cross-cultural, developmental, animal and



neuroscientific studies that support the existence of innate principles of moral thought.

In other regards, however, language seems very different from morality. For one thing, linguistic knowledge is distinct from emotion. You might be disgusted or outraged by what somebody says, but the principles that make sense of sentences are themselves entirely cold-blooded. Your eyes do not well with tears as you unconsciously determine the structural geometry of a verb phrase. By contrast — and Hauser wrestles with this throughout *Moral Minds* — even those who accept that some moral capacity is innate often see it as inextricably linked to emotion. Perhaps the universal core of morality is a set of emotional responses — disgust, shame, sympathy, guilt and so on — that are triggered by certain situations. This hypothesis is supported by clear demonstrations that, at least in some circumstances, emotion precedes intuition. The psychologist Jonathan Haidt of the University of Virginia in Charlottesville has found, for instance, that we are sometimes first viscerally shocked by

a scene ("Omigod, he's having sex with a chicken!") and then converge upon a moral judgement ("There should be a law against that.").

A different concern is that languages are combinatorial symbolic systems. An English speaker, for example, knows perhaps hundreds of thousands of words, and also knows principles of syntax that dictate how these words combine with one another to form sentences. There are other combinatorial systems in human cognition, such as number and music, but it's not clear that morality is one of them. Even if it is distinct from emotion, moral knowledge might be better characterized as a small list of evolved rules, perhaps simple (such as a default prohibition against intentional harm), perhaps complex (such as some version of the doctrine of double effect), but still very different in character from linguistic knowledge.

This combinatorial issue becomes relevant when it comes to differences in morality between people. The approach developed by Chomsky explains differences in human languages in terms of the parametric variation of universal principles. All languages have verb phrases, for instance, but in some of them, such as English, the object follows the verb, whereas in other languages the object goes first. Hauser makes an excellent case that the variation in moral systems is constrained in interesting ways, but he provides no evidence for parametric variation of the linguistic sort. Instead, as the cultural anthropologist Richard Shweder of the University of Chicago, Illinois, and others have argued, you get a difference in emphasis: all cultures value both purity and fairness, for instance, but some emphasize the former and others the latter. In addition, there can be dramatic variation in morality within an individual culture. As Hauser notes, for example, the culture of honour in countries such as Pakistan leads to a shockingly high number of 'honour killings' every year. However, that does not imply that all native Pakistanis agree

ILLUSTRATIONS BY JOE MAGEE

on the moral permissibility of honour killings as they do on the linguistic grammaticality of well-formed Urdu sentences.

These differences call into question how much insight the research programme developed by linguists can provide into morality. For instance, some arguments for linguistic innateness are based on assumptions about the generative nature of language; these might not export well to the moral domain. In fairness, though, it might be that nobody has found

these sorts of deep parallels because, before Hauser, nobody had really looked. Moreover, even if morality lacks certain interesting features of language, the very idea of an innate moral faculty is well worth investigating. Linguists have been exploring this idea for more than 50 years; in the study of morality, we are just getting started.

Paul Bloom and Izzat Jarudi are in the Department of Psychology, Yale University, New Haven, Connecticut 06520-8205, USA.

is done well, although the tone is sometimes gratingly folksy, and there is some unevenness in the level of knowledge assumed (nanometres and natural logarithms are used without definition, for example). The explanations are clear and easy to follow, and the level of factual accuracy is very high, although effective population size is incorrectly used instead of population size in the book's only equation.

Carroll also turns his attention to irrational views on a variety of topics, including the genetics of Trofim Lysenko, chiropractic 'medicine', and opposition to vaccination. This paves the way for a discussion of disbelief in evolution, illustrated with some remarkable quotations from US anti-evolutionists. I am stunned by the ability of people to accept the absurd and discount the rational. Carroll seems confident that the mainstream religions do not reject evolution, but he fails to mention Islam, for which this is far from clear. He quotes with approval a statement by Pope John Paul II on evolution, but does not refer to Cardinal Christoph Schönborn's anti-darwinian diatribe in *The New York Times* on 7 July 2005. Carroll's final chapter warns of the dangers arising from climate change and the overexploitation of natural resources, where scientific advice is often ignored by governments because of economic considerations.

As Carroll rightly concludes: "Understanding and accepting evolution is a matter of adhering to the scientific process." It is extremely important

for scientists to fight the idea that evolution can be separated from the rest of science simply because of its unpalatable implications for traditional religious accounts of human origins. We have to insist on the fact that scientific conclusions are based purely on the study of nature, without reference to religious or other authority.

Carroll's book will certainly help the public to understand evolution

more clearly, but rational arguments are unfortunately unlikely to persuade those whose beliefs come from prior authority, whether represented by a fatwa, an encyclical, or the first chapter of Genesis. Wider public understanding of the nature of scientific evidence in general is urgently needed.

Brian Charlesworth is at the Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 2BR, UK.

Evidence for evolution

The Making of the Fittest: DNA and the Ultimate Forensic Record of Evolution
by Sean B. Carroll

W. W. Norton: 2006. 301 pp. \$25.95

Brian Charlesworth

Sean Carroll begins his excellent book *The Making of the Fittest* by pointing out that about 50% of the American public doubt the truth of darwinian evolution, yet accept other aspects of biological science, such as the use of DNA in forensics. His aim is to use evidence from modern research on DNA to convince the general reader of "the case for biological evolution as the basis for life's diversity, beyond any reasonable doubt". He deliberately does not introduce any of the standard evidence for evolution as a historical process, and only briefly describes the use of DNA sequence data to reconstruct phylogenies. Instead, he uses a series of examples intended to provide compelling evidence for the basic processes involved in evolutionary change, with particular emphasis on mutation and natural selection.

The first example involves the extraordinary icefish of the Antarctic, which have lost their red blood cells and allowed their haemoglobin genes to decay. Carroll shows how this can be interpreted as an adaptive response to life in very cold water, with its high solubility of oxygen in water and high blood viscosity. There follows an exposition of Darwin's ideas about natural selection. Next comes a good but simplified account of the standard theory of selection on single mutations that enhance fitness,

showing that these can spread to high frequencies in times that are trivially short compared with the geological record. Carroll then applies these results to the way that a species of desert-living mouse has adapted to spatial differences in the colour of the ground, right down to the mutational changes that lead to darker coat colour.

Carroll illustrates the power of natural selection to generate evolutionary novelties by discussing duplications of the opsin genes involved in vertebrates' colour vision. Examples of genes and genomes that decay when their function is no longer maintained by natural selection illustrate how "selection acts

only in the present, and not as an engineer or designer".

A favourite ploy of creationists is to accept the possibility of small-scale evolutionary change by darwinian means, but to deny that this has any relevance to the evolution of complex structures or new species. Carroll does not discuss how new species evolve, but he examines the problem of complexity using Darwin's example of the eye, bringing in recent results on the molecular basis of eye formation. All this



Battlefield between the ears

Mind Wars: Brain Research and National Defense

by Jonathan D. Moreno

Dana Press: 2006. 225 pp. \$23.95

Charles Jennings

On an evening in October 2002, a group of armed Chechen separatists overran a Moscow theatre, taking the audience hostage and rigging the building with explosives. After a short stand-off, Russian special forces flooded the building with an aerosol based on the potent synthetic opioid fentanyl and then stormed it. The narcotic was effective in incapacitating the hostage-takers, many of whom were executed on the spot. In the ensuing chaos, however, 117 hostages also died from fentanyl poisoning. They were victims not only of terrorism but also of poor planning: if the opioid antagonist naloxone had been available to rescuers, many of these deaths could have been avoided.

Welcome to the world of *Mind Wars* and the military application of neuroscience, which is the subject of this fascinating and sometimes unsettling book. As the author Jonathan Moreno reveals, the US military has a long-standing interest in brain research and, as scientific understanding continues to advance, so does its appeal to the national security establishment. The Department of Defense conducts much of its research in secret, and some of it would probably fare poorly in open peer review — for example, the military continued to fund psychic research until 1995 — but with an annual research and development budget of at least \$68 billion, it can presumably afford to leave no stone unturned. Partly because its activities are more visible, Moreno focuses especially on the Defense Advanced Research Projects Agency (DARPA), which supports unclassified academic research with potential military applications. DARPA has a distinguished record of supporting innovation, including the Internet, so its involvement in brain research must be taken seriously.

A key aim for DARPA is to enhance the mental capacity of human warriors, whose brainpower is an ever-increasing bottleneck in an age of high-tech warfare. Brain-monitoring devices could help manage the flow of information delivered to soldiers in combat, and their mental performance might be further enhanced through drugs — an anti-sleep pill, for instance, is a target for the military. More speculatively, prosthetic devices might even allow human operators to control machines directly from the brain, with greater speed and bandwidth than can be achieved through manual controls. Brain-machine interfaces are still in their infancy but are developing fast, thanks in part to DARPA funding most of the leading academic labs in the field.

As the Moscow siege dramatically illustrated,

brain science can also be harnessed for direct offensive means — targeting the brains of one's opponents. The US military has a long-standing interest in (supposedly) non-lethal targeting methods, including drugs as well as less conventional approaches such as loud noise and noxious smells (the most potent of which is said to be 'US Government Standard Bathroom Odor').

Another potential application is the field of interrogation. Future advances might allow us to determine through brain scanning whether someone is telling the truth, and perhaps even to devise methods for reducing their ability to lie. The military began investigating candidate 'truth drugs' (including LSD) in the early 1950s, and although the goal remains elusive, it seems sensible to ask what possibilities exist and under what circumstances it would be acceptable to use them.

Moreno is a bioethicist, and a focus on ethical questions runs through the book. Although I found some of his concerns overstated — he worries, for instance, that soldiers would be depersonalized through brain implants, or that advances in neurogenetics will tempt us to breed genetically altered warriors who lack fear — others seem well founded. If military research leads to effective mind-reading techniques, for example, there will inevitably be pressure to extend these to the civil judicial system, which will raise urgent questions about cognitive liberty.

One weakness of the book is that Moreno's treatment of technical issues is sometimes superficial. Some of the ideas discussed here — such as brain scanning at a distance, or memory augmentation through hippocampal implants — fall close to the fine line separating the visionary from the crackpot, and a more critical examination of this border territory would have been welcome. Moreno recognizes the outright nonsense as such, but an over-reliance on popular news stories rather than technical sources sometimes leads him to give outlandish ideas more credence than they deserve.

The book is a fascinating read despite these reservations, but it still left me wondering, is this stuff for real? Moreno implies it is when

he writes: "the powers that can...establish a 'neurotechnology gap' between themselves and their adversaries will establish tactical and strategic advantages that can render them dominant in the twenty-first century." I'm no expert on warfare, but this strikes me as hyperbole. As a reality check, I recommend reading *Fiasco* (Penguin, 2006), Thomas Ricks' devastating critique of the United States' botched war in

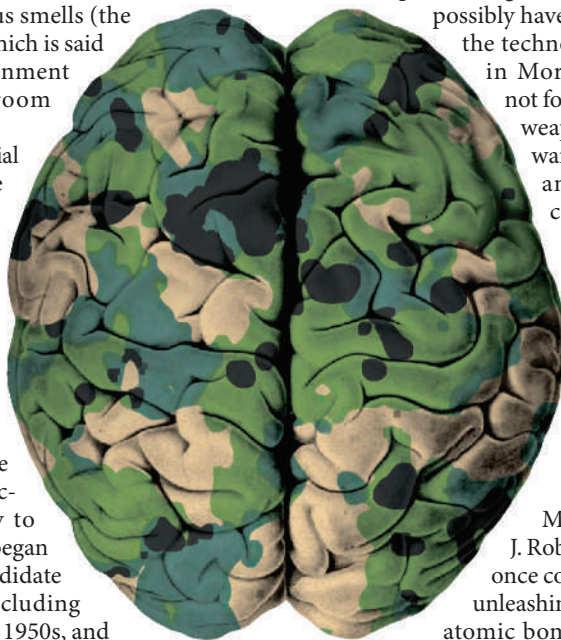
Iraq, and asking how the outcome could possibly have been changed by the technologies envisaged in Moreno's book. It is not for lack of high-tech weapons that the Iraq war is going so badly, and given America's overwhelming superiority in firepower, it is not obvious how a new generation of brain-based weapons would represent more than an incremental gain.

The leader of the Manhattan project, J. Robert Oppenheimer, once commented that, by unleashing the power of the atomic bomb, the physicists involved had "known sin". The

activities described in *Mind Wars* may seem like a harmless flirtation by Oppenheimer's standard. Yet any academic involvement in military research presents an ethical dilemma, and Moreno's exploration of this theme is one of the most interesting aspects of his book. He is no knee-jerk pacifist: he accepts that military force is sometimes necessary and argues convincingly that contact between military and civilian research is healthier than the alternative of total secrecy. He also acknowledges the 'dual-use' argument that many DARPA-funded programmes have clear civilian pay-offs. Yet by taking military funding, he says, researchers are in some sense accomplices to the perpetuation of what he calls a "national security state", a posture of open-ended militarization supported by a vast budget that, in the view of many critics, bears little relation to the actual threats confronting the United States.

When interviewing neuroscientists for his book, Moreno found many of them reluctant to discuss the issues raised by their involvement with defence-related research. *Mind Wars*, the first systematic treatment of the topic, should help bring these questions into the open. ■

Charles Jennings is a consultant based in Concord, Massachusetts. He is a former editor with the *Nature* journals and a former executive director of the Harvard Stem Cell Institute. Supplementary links for this review are available at www.connotea.org/user/charlesj/tag/'Mind Wars'.



Experimental theatre

Science on Stage: From *Doctor Faustus* to *Copenhagen*

by Kirsten Shepherd-Barr

Princeton University Press: 2006. 271 pp.
\$29.95, £18.95

Stuart Firestein

Copenhagen (Michael Frayn, 1998), *Proof* (David Auburn, 2001), *Wit* (Margaret Edson, 1995), *Arcadia* (Tom Stoppard, 1994) and *A Beautiful Mind* (Ron Howard, 2001): all recent winners of a Pulitzer, Tony, Olivier or Oscar, and all dramas about science and scientists. Along with the recent proliferation of television shows that feature science (the ubiquitous forensic-investigation series, for instance), these examples seem to give the lie to the Janus-faced cultural split between the humanities and sciences postulated famously by C. P. Snow in 1959. In the theatre, science is current, popular and topical. In the United States, for example, the Alfred P. Sloan Foundation, in a programme devoted to increasing public understanding of science and technology, devotes significant funds to encouraging artists and playwrights to create new works in the theatre using science themes, including financing productions as part of the First Light Festival in New York.

This intersection of research and performance is chronicled in *Science on Stage* by Kirsten Shepherd-Barr of the University of Birmingham, UK (I should disclose a connection here: Shepherd-Barr is the daughter of the Yale neuroscientist Gordon Shepherd, my postdoctoral mentor). As she points out, it is not an entirely new phenomenon: an eye-opening appendix lists 122 plays that make central use of scientific subjects, beginning with Marlowe's *Doctor Faustus* (published posthumously in 1604) and Jonson's *The Alchemist* (1610), and covering a further four centuries of theatrical literature and performance. Shepherd-Barr describes, analyses and interprets a host of theatrical scripts and performances with science as their themes. She provides important historical context and lots of interesting insights, especially regarding the most recent plays.

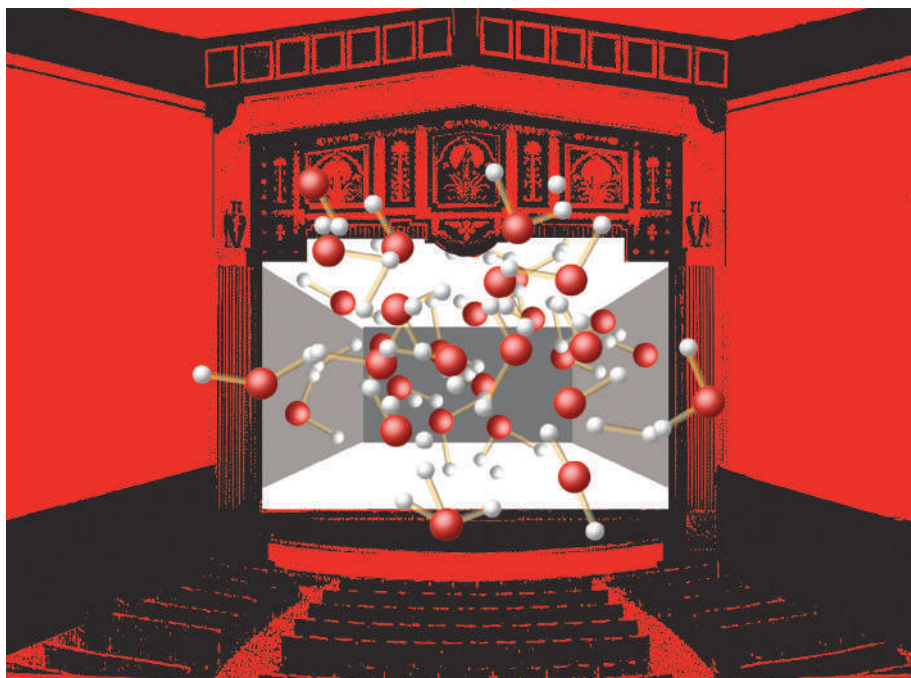
Several interesting issues come up in this book. Science is becoming more public as it insinuates itself ever further into the lives of individuals and into decisions that we must make as a society. The theatre is perhaps the most public of the arts, the prototypical forum for the portrayal of ideas and the working-out of conflicts. It is the ideal forum for a meeting of the two cultures — a venue for the transformation of science, often deemed obscure and aloof, into personal drama. There are many commonalities: both science and the theatre are engaged in describing the world, and science, according to the dramatist Peter Brook, is

the new mythology, the place to which people turn for answers. Science explores many of the eternal mysteries that have been the subject of theatre since the first performances: how we got here, what we're doing here, and where we're going.

Science themes appear in other art forms as well, but Shepherd-Barr points out that the theatre has been especially hospitable to science; in particular, more so than film, which as a medium can't seem to avoid turning science into fantasy. She maintains that for creative,

of the productions to shed light on the difficulties and the opportunities that arise from turning science into theatre. *Copenhagen*, for instance, made some unprecedented demands on the actors with regard to the amount and difficulty of memorization required. Shepherd-Barr also discusses the particular problems that playwrights using science have in presenting exposition, and the inevitable tension between historical and scientific accuracy and dramatic demands.

Also considered here are scientists turned playwrights, the leading example being Carl Djerassi, emeritus professor of chemistry at Stanford University, California. His effort with Nobel-prizewinning chemist Roald Hoffman,



thoughtful writers, science opens up new territories of rich material, beyond the stale melodramas of dysfunctional families, interpersonal soap-opera plots and psychological mysteries. In her most insightful comments she shows how the best science plays, *Copenhagen* or *Arcadia* for example, use the science they deal with — atomic structure and entropy respectively — within the structure of the play.

The book begins with an overview and an essay that tackles the interesting question of the basis for the current appeal of science plays. The middle chapters document plays in the various sciences, touching on physics, evolution, medicine and mathematics. Within them there are detailed analyses of *Copenhagen* and *Arcadia* as prime examples of successful plays that are both theatrically engaging and scientifically satisfying. These are especially enjoyable chapters and, although familiarity with the plays is not required for their appreciation, for anyone who has seen them it will be fun to have the performances brought back to life. These are not simply rehashes of reviews; instead, Shepherd-Barr uses the scripts and features

Oxygen (2001), combines a wealth of scientific information in an engaging plot surrounding the discovery of the gas.

A final chapter looks at unconventional theatrical presentations, in particular collaborations involving both a scientist and a dramatist or director. Examples include innovative productions by cosmologist John Barrow of Cambridge University, UK, and director Luca Ronconi (*Infinites*, 2002; see *Nature* **416**, 585; 2002); dramatist Jean-François Peyret and neurobiologist Alain Prochiantz of the National Centre for Scientific Research (CNRS) in Paris, France (*Darwin Variations*, 2004; see *Nature* **432**, 445; 2004); and Peter Brook and neurologist Oliver Sacks (*The Man Who*, 1993). This is an especially lively portion of the book and leaves one, rightly, with the exciting idea that these productions are just the beginning. Appropriately, the embracing of scientific themes has motivated experimentation with theatrical form and content. The descriptions of these new performance pieces make one sorry to have missed any of them. *Infinites* and *Darwin Variations*, unaccountably, have

never been translated or performed in English.

Who will read this book? Who should read it? Scientists, playwrights, directors, humanists. It is unquestionably a scholarly work, comprehensive, utilizing many sources, thoroughly referenced and heavily annotated. Parts of it read as if they are composed of essays that could stand alone, making it occasionally repetitive. On the other hand, it is a surprisingly breezy and entertaining read.

One could hope that this book might insti-

gate university courses, perhaps taught jointly by professors in science and theatre departments. Such a course could be equally engaging to students in physics, chemistry, biology, mathematics, literature, drama, history — whoever heard of such a thing? As a working scientist I am left with only one question: why aren't there any science comedies? ■

Stuart Firestein is in the Department of Biological Sciences, Columbia University, New York, New York 10027, USA.

Sermons and straw men

The God Delusion

by Richard Dawkins

Bantam/Houghton Mifflin: 2006. 416 pp
£20/\$27

Lawrence M. Krauss

Early in *The God Delusion*, Richard Dawkins quotes Thomas Jefferson's statement: "I am satisfied, and sufficiently occupied with the things which are, without tormenting or troubling myself about those which may indeed be, but of which I have no evidence."

This eminently scientific sentiment was echoed two centuries later by the physicist Steven Weinberg, who, like Dawkins, is an outspoken critic of religion, but who has nevertheless suggested that most scientists simply don't spend enough time even thinking about God or religion to merit the label atheist.

But Richard Dawkins is a man on a mission. This new book is the culmination of his recent campaign, which has included a two-part television documentary, aired in Britain, to help humanity rid itself of what he surely views as one of its most vile creations: God.

Before I proceed further, I should, in the interests of full disclosure, confess that I have written exactly one fan letter that I can remember to an author. That letter was written to Richard Dawkins after the publication of a small book, *River Out of Eden* (Weidenfeld & Nicolson, 1995), which I felt was perhaps the most concise and cogent science writing, as well as the clearest discussion of the nature of evolution, I had yet read.

I wish that Dawkins, who has a gift for making science — in particular, evolutionary biology — both exciting and understandable to a broad audience, had continued to play to his strengths, which are desperately needed now more than ever as we confront growing attacks on the teaching of evolution, not just in the United States but in the UK and Europe.

Dawkins the preacher is less seductive. And make no mistake: this book is, for the most part, a well-referenced sermon. I just have no idea who the intended parishioners might be. In his preface, Dawkins claims he hopes to reach religious people who might have misgivings, either about the teachings of their

faith or about the negative impact of religion in the modern world. For these people, Dawkins wants to demonstrate that atheism is "something to stand tall and be proud of".

I found this slightly puzzling. I don't believe in Santa Claus, but I am not particularly proud of it. Indeed, I am rarely, if ever, proud of not believing in things. More generally, I think the strategy of focusing on telling people what not to believe is less compelling than positively demonstrating how the wonders of nature can suggest a world without God that is nevertheless both complete and wonderful — an argument that Dawkins reserves for the final few pages of the book. And while there is a lot to complain about in the ubiquitous facile piety so prevalent today, complaining can nevertheless start to get tiresome. Carl Sagan's *The Demon-Haunted World* (Headline, 1996) likewise too often mirrored Sagan's frustration at all those who over many years have continued to confront him with their superstitions, but it also conveyed his sense of awe and wonder about nature in a way that Dawkins elsewhere has so craftily displayed.

A less sympathetic reader than the author's wife (who apparently read the entire manuscript aloud to Dawkins for him to review) might have provided a more useful foil. Several indulgences detract from the flow, but more importantly,

I was struck at how Dawkins' presentation, particularly in the early chapters where he builds his case against God, might offend those who, like myself, are quite sympathetic to his central thesis. I suspect that few thinking people of faith are unaware of the remarkable evil that has been done in the name of God, or the possibility that, although most cultures worship some god, this could be a mere reflection of the workings of the human brain rather than definitive evidence for God's reality. Yet Dawkins seems to suggest early on that even agnostics might never confront these issues and that he needs to "raise their consciousness", as he puts it. At the very least I find it doubtful that constantly questioning the intelligence of 'true believers' will be helpful in inducing any such reader to accept Dawkins' strongly argued thesis that both God and religion are nonsensical and harmful.

While I usually tend to begin a review with praise and end with reservations, the reverse order here reflects the progression of my own reading of *The God Delusion*. There are gems in the book, as one might expect from a writer as powerful as Dawkins, but most of them are in the second half of the volume. He ends the first half with what I found to be a less than compelling probabilistic argument against God. (Incidentally,

I couldn't help wondering, somewhat facetiously, when

Dawkins used an anthropic argument from cosmology to argue against God, that, although indeed only very rare universes may harbour life, if an infinite number of universes exist, could not at least one then harbour what might pass for a divine being?)

But after this Dawkins proceeds to brilliantly review the

roots of modern morality and the changing moral 'zeitgeist', as he calls it. With authority and wit, he marvelously dissects the absurdity, hypocrisy and selectivity that is inherent in so much of modern biblical morality. Perhaps there can be no higher praise than to say that I am certain I will remember



and borrow many examples from this book in my own future discussions.

Finally, his treatment of religion and childhood is, I believe, precisely accurate. We do our children a great disservice (which Dawkins goes so far as to call abuse) by forcing religion upon them at an age where they are far too young to digest the deep and subtle issues associated with the possibility of divine purpose. In doing so, we encourage them to rely on potentially destructive emotions rather than to use their brains. And pondering how to build a world of adults that might result from the latter

rather than the former is surely what motivates Dawkins here, even if his approach might also provoke some religious fundamentalists to harden their unfortunate belief that knowledge and reason are dangerous things, to be avoided at all costs.

Lawrence M. Krauss is Ambrose Swasey Professor, and director of the Center for Education and Research in Cosmology and Astrophysics, at Case Western Reserve University, 10900 Euclid Avenue, Cleveland, Ohio 44106-7079, USA. His most recent book is *Hiding in the Mirror*.

sparse resources) as well as the ups (discovery of totally new phenomena, fast friendships) of field work. Key incidents, such as the Gombe kidnapping, in which four students were held for ransom by Congolese rebels, are revealing; for instance, for weeks in captivity, they had but one English-language book, Niko Tinbergen's *The Herring Gull's World* (Collins, 1953).

Heroes and villains abound. A series of older male mentors, such as Robert Hinde at Cambridge and David Hamburg at Stanford University, California, were instrumental in guiding the neophyte Goodall. Others were less helpful, such as the zoologist Solly Zuckerman, who, in chairing the symposium at which Goodall first presented her results, dismissed them as fantasy. Lesser egos would have been crushed. The National Geographic Society was both saviour and censor, providing vital funding in the early days, but repeatedly restricting her efforts to present her findings. Easily the most influential figure, however, was Leakey, who although crucial to Goodall's emergence, also complicated her life on personal and professional fronts. Having offered her the Gombe project, he then offered it to another attractive young woman, and kept Goodall waiting for months, until the other turned it down.

Peterson gives the reader useful end-notes but a limited list of references: this is not the place to learn about primatology. The photographs are disappointing, all being in black and white, and mostly snapshots.

The book could have been called 'The Many Lives of Jane Goodall', given her several career transitions: from chimpanzee expert, to distinguished scientist, to committed conservationist, to animal-welfare activist, to global campaigner for the environment, to eminent 'world citizen' in the causes of youth and peace. She would be famous for any of these incarnations, although Peterson is a bit over the top in

terming her the most visible scientist of the twentieth century (cf. A. Einstein) and the best-known female scientist in history (cf. M. Curie). Regardless of this, Peterson's volume is the one for which Goodall's devotees have been waiting, and it will deservedly enjoy much success.

W. C. McGrew is at the Leverhulme Centre for Human Evolutionary Studies, Department of Biological Anthropology, University of Cambridge, Cambridge CB2 3DZ, UK.

The hero of Gombe

Jane Goodall: The Woman Who Redefined Man

by Dale Peterson

Houghton Mifflin: 2006. 672 pp. \$35

W. C. McGrew

In this age of hyperbolic subtitles — such as Frans de Waal's *Our Inner Ape: A Leading Primateologist Explains Why We Are Who We Are* (Riverhead, 2005), or Richard Dawkins: *How a Scientist Changed the Way We Think* (Oxford University Press, 2006) — comes another with a major claim. Dale Peterson has produced the long-awaited, definitive biography of (arguably) the world's leading figure in primatology.

He is surely well qualified to do so, not only as an award-winning author, but also as the editor of two previously published volumes of Goodall's letters, as well as the co-author with her of an influential book that crossed disciplinary lines (*Visions of Caliban*, Houghton Mifflin, 1993). But might he be too close to his subject to be able to give us an objective account? Might his ambitious subtitle be exaggerated? Happily, the answer to both questions is no.

Goodall is first and foremost associated with wild chimpanzees — more particularly, with one small population of these African apes living on the eastern shores of Lake Tanganyika, in western Tanzania. Her study of the chimpanzees of Gombe, begun in 1960, continues to this day, and has spawned many books, films and articles, both scientific and popular. Indeed, our view of that species, which (along with the bonobo) is humanity's nearest living relation, was shaped, at least in the West, by Goodall's early researches, as reported widely by the National Geographical Society. Few of us have not seen the evocative images of the slim, pony-tailed young woman and her hairy subjects in the Edenic surroundings of Gombe.

But what of the subtitle's claim, that she has redefined our species, *Homo sapiens*? This originates in the response of Louis Leakey, the flamboyant archaeologist and mentor of Goodall, to her unexpected finding that the apes not only used tools but also made them, as

part of their extractive foraging. He stated, with characteristic panache, that now we must redefine 'tool', redefine 'man', or accept chimpanzees as humans. This blurring of long-established boundaries was further advanced by Goodall's observations of chimpanzees doing other human-like things, such as hunting, cannibalism, warfare, adoption and ritual-like displays.

The book is comprehensive, following Goodall from her childhood interest in animals through her somewhat chequered young adulthood as debutante and waitress. The breakthrough came with a self-financed trip to Kenya, where she met Leakey, who set her on the route to primatology, although she had no prior training nor higher education. Her aptitude showed quickly as she met chimpanzees on her first day in the field and discovered tool use and meat eating within the first few months. These successes led to her acceptance at the University of Cambridge, UK, to do a doctorate in ethology, despite not having a first degree. From that point onwards, she never looked back.

Peterson writes vividly. Descriptions of early days at Gombe, based on excerpts from Goodall's field notes and letters home, come alive. Days in the forest or on the savanna capture the downs (ill health, frustrations with



The theorist

Francis Crick: Discoverer of the Genetic Code

by Matt Ridley

HarperCollins: 2006. 213 pp. £12.95, \$19.95

Horace Freeland Judson

Any biographer of Francis Crick faces a difficult problem: he was the greatest biologist of the latter half of the twentieth century, yet his life was curiously one-dimensional. It was intensely concentrated but narrowly focused. By far the most important part of his life was his science, and because of his pre-eminence and unique role, his life was the life of his science in his day. And what a day it was — the golden age of molecular biology.

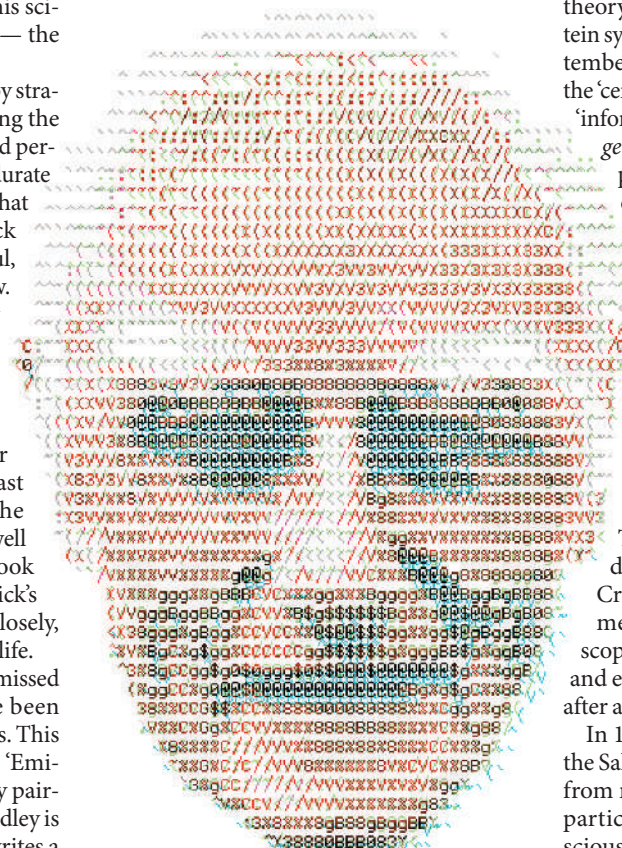
A biography will therefore be shaped by strategic choices: how far to go in addressing the other dominant aspects of Crick's life and personality — his vivid sexuality and his obdurate atheism — and, more importantly, at what level of detail to treat the science. Crick was erotically hyperactive and successful, although not predatory as far as I know. Matt Ridley does write of this, lightly and tactfully — an acknowledgement rather than an examination. That's fine. The atheism was rooted and militant. Ridley introduces it where it determined Crick's choice of problems: molecular biology in the late 1940s to dispel the last traces of vitalism, neurobiology in the 1970s to demystify consciousness. All well and good. But I came away from the book regretting that Ridley did not engage Crick's science and his intellectual style more closely, because here lies the richness of Crick's life.

As Crick's first biographer, Ridley has missed a unique opportunity. He could have been definitive, pre-empting later biographers. This book is one of a series, under the rubric 'Eminent Lives', intended to be short, snappy pairings of top people with strong writers. Ridley is an established science journalist who writes a pleasant, ambling prose. But he has not reached to the well-springs of Crick's singular role.

The subtitle is symptomatic, for Crick was both less and far more than "discoverer of the genetic code". Breaking the code was a paradigmatic example of collective-competitive effort. The idea that there must be a code had been put forward in 1944 by the physicist Erwin Schrödinger. From the moment in the spring of 1953 when Crick and James Watson announced the structure of DNA, the need to determine how the structure carries the code was obvious. That summer, the eccentric physicist George Gamow was the first to propose an actual mechanism, wildly wrong but forcing the question. The task itself — to identify which of the 64 triplet codons of the four bases in DNA specify which of 20-odd amino acids — took several years and at least

half a dozen scientists and laboratories to solve. Crick himself did very little of the laboratory work. Rather, he was the explicator, the arbiter, the taskmaster.

Crick was above all the theorist — and that's the unifying thread. Not just with the coding problem but throughout his career, Crick was the one who put other scientists' thoughts in order. He soaked up data, mostly other people's, but saw beyond the data to their meaning, their shape, their implications. He found principles, and did it with a preternatural clarity of mind and in inimitable style.



In the interplay with Watson that led to the structure of DNA, it is clear that while Watson had the immediate visual insights, Crick framed their approach in general terms. Ridley, by necessity, tells the old familiar story of the discovery, but he fails to capture the nature of the intellectual interaction — or the psychological dimension, which was driven by Watson's intense admiration of Crick's prowling, relentless intelligence and by his envy of Crick's success with girls.

Ridley also fails to reach the origins of the historic conflict between the other scientists working for the structure, Rosalind Franklin and Maurice Wilkins at King's College London. Wilkins had been getting important new X-ray-diffraction patterns from DNA. Franklin, an X-ray crystallographer, was recruited in

the spring of 1950 (not in December, as Ridley writes). She joined the lab in January 1951, and Ridley makes no mention of a meeting she had then with the physicist Alexander Stokes and a graduate student, Raymond Gosling, at which John Randall, the lab's director, gave her a supply of the best DNA they had and appointed Gosling her assistant. Crucially, Wilkins was away on vacation. Franklin had every reason to think the DNA was exclusively hers. When Wilkins returned and expected to collaborate with her, she shut him out. He grumbled about her to Crick and Watson, and in February 1953 he notoriously showed Watson an X-ray diagram she had obtained — which they interpreted as she had failed to do.

Crick's boldest, most enduring exercise in theory came in a remarkable paper, "On protein synthesis", that he read at a meeting in September 1957. It culminated in what he called the 'central dogma' of molecular biology: "Once 'information' has passed into protein it cannot get out again." Information here meant "the precise determination of sequence", either of bases in nucleic acids or of amino acids in proteins. This was a characteristically practical assertion. At a time when so much was still to be learned, it ruled out great swaths of speculation. Yet it has great philosophical force, for it is a radical statement of the physiological reason why the inheritance of acquired traits cannot occur.

The paper forever altered the fundamental logic of biology. And it epitomizes Crick's essential intellectual style. Throughout the book, Ridley's account, despite ticking off the wide range of Crick's scientific interests and accomplishments, seems to convey only fuzzily the scope, the rigour, and beyond that the clarity and energy that Crick displayed — the basis, after all, of his dominance of the field.

In 1976, Crick moved from Cambridge to the Salk Institute in La Jolla, California — and from molecular biology to neurobiology, in particular the neural, cellular basis of consciousness. He winnowed the literature, then with new colleagues put forth a succession of ideas and strictures. As he himself recognized, he did not get far. Yet, as Ridley writes, Crick's interest, his unflagging skill at incisive generalization, and his reputation made the physiology of consciousness respectable. His style persisted.

Ridley's book is marred by factual errors, which perhaps ought to be beneath the attention of a brief review, yet they occur in such a swarm that they sap the reader's confidence. For example, the first experimental sighting of what came to be recognized as messenger RNA was made not by two Russian scientists, as Ridley writes, but by Elliot Volkin and Lazarus Astrachan, both American born, who worked at the Oak Ridge National Laboratory in Tennessee. Again, writing of the French molecular biologist Jacques Monod, Ridley lists his

remarkable qualities: sailor, cellist, and so on, and includes "communist", without explanation. This is a reprehensible distortion. Monod joined the French communist party during the war for just one reason, he told me: it was the only way he could have any influence on the strategy and tactics of the armed resistance, in which he then rose to chief of operations for

all of France. After the war he quietly dropped out, until in 1948 he publicly denounced the fraudulent claims of Stalin's geneticist and science chief, Trofim Lysenko — and with terrific effect.

Horace Freeland Judson is at 807 West University Parkway, Baltimore, Maryland 21210, USA.

of the bayesian approach. Pragmatically, the two interpretations should be considered as complementary, rather than competing. Undoubtedly, bayesian inference has profound implications as it concerns critical thinking and may lead to a deeper understanding of the scientific method. However, the ability to develop a new hypothesis or theory relies in the end on scientists' creativity, something that can hardly be delimited within a methodological framework.

I have a minor observation. Randomness, which is intrinsically related to probability, deserves a more detailed discussion than is given here. In particular, the Monte Carlo method is mentioned only briefly, despite its importance and wide use in virtually every branch of science.

Altogether, the book provides a truly enjoyable overview of the role of probability in science, as well as in everyday life. It is aimed essentially at non-specialist readers, but even those who are familiar with its contents will enjoy the stimulating presentation. Finally,

Coping with uncertainty

From Cosmos to Chaos: The Science of Unpredictability

by Peter Coles

Oxford University Press: 2006. 224 pp.
£25, \$44.50

Gianpietro Malessio

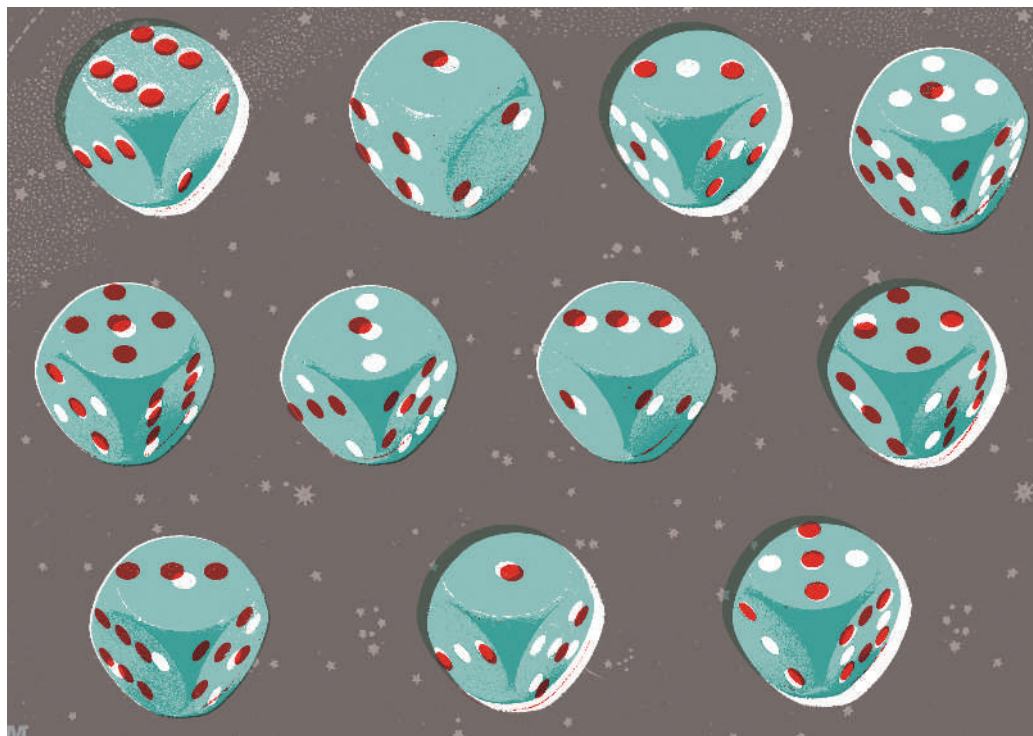
If the process through which civilization evolved were condensed into a few words, these could be: posing questions, finding answers. As those with children know well, wondering about the world is intrinsic to human nature. This attitude is generally lost as we grow up (life is a tough business), but a good scientist retains it as an adult. In principle, finding answers is the ultimate goal of a scientist. However, scientific answers have a peculiar character: they do not represent absolute truth acquired once and for all, but only temporary beliefs subject to continuous, critical revision. Scientific knowledge is anti-dogmatic: it represents the most reliable conclusion that can be drawn on the basis of the available data. There are no certainties in the realm of science. This may seem profoundly disappointing to the general public, but it should not. Science is a way of thinking much more than it is a body of knowledge. Its true achievement is not providing answers, but rather a method through which answers can be sought.

A basic instrument of this method is represented by probability theory. As Peter Coles brilliantly illustrates in his engaging new book, *From Cosmos to Chaos*, probability theory makes it possible to extract meaningful information from a sea of noisy data, to deal with intrinsically probabilistic processes, and to infer conclusions on the basis of incomplete information.

After explaining the basics of probability theory, Coles offers a ride across virtually the entire spectrum of the physical sciences, including thermodynamics, cosmology, chaos theory and quantum mechanics, as well as stimulating issues such as the anthropic principle and life on other worlds. Each topic is considered from the specific perspective of how it involves, and is related to, the concept of probability. Coles also examines the impact of probability on many aspects of everyday life. He shows us how the decisions of law courts and medical tests, as well as more mundane activities such as gambling strategies, may

suffer deeply from an incorrect use of statistics.

To illustrate how probability theory works, Coles presents some intriguing and often surprising examples, encouraging readers to think for themselves and providing them with material to entertain and amaze their friends. I particularly appreciated the historical notes and lively anecdotes that highlight the famous and less well-known personalities who contributed to the development of the statistical



approach to science. Coles points out that most of these were astronomers. Historically, men of science generally applied their talents to a range of disciplines, so most of the people cited in the book might also be considered as mathematicians and physicists.

A central chapter of the book is devoted to the controversy between the frequentist and the bayesian interpretations of probability. The former interprets probability as a frequency of occurrence in repeated experiments, whereas in the latter probability measures the degree of belief an individual has in an uncertain proposition. As Coles openly declares, one of the book's objectives is to argue in favour

I fully endorse Coles' passionate plea at the end of the book for a better integration of science and society. To face challenges such as large-scale pollution, the energy crisis or food shortages, we will need all the instruments that science puts at our disposal. Perhaps an even greater threat is represented by fundamentalisms of any kind (religious, political, economical) that aim to impose their dogmas on the whole of society. Science, or rather, the *forma mentis* it gives us, may offer the best antidote for defying them.

Gianpietro Malessio is in the Department of Physics, University of Messina, 98166 Messina, Italy.

NEWS & VIEWS



K. WOTHE/PHOTOLIBRARY.COM

GENOMICS

How to make a social insect

Edward O. Wilson

The profound biological changes that lofted the honeybee to an advanced state of social organization are reflected in its newly sequenced genome. The species can now be studied all the way from molecule to colony.

The transformation of an insect species from a solitary lifestyle to advanced colonial existence requires alterations in every system of the body, coupled with sufficient plasticity in the traits prescribed by the genes to generate strong differences among the adult castes. A picture of this revolution at the genomic level is published in this issue (page 931)¹.

If Earth's social organisms are scored by complexity of communication, division of labour and intensity of group integration, three pinnacles of evolution stand out: humanity, the jellyfish-like siphonophores, and a select assemblage of social insect species^{2,3}. The honeybee, *Apis mellifera*, is a member of this insect group — which also includes the spectacular leaf-cutting ants, army ants and macrotermitine mound-building termites — and no one can deny that it belongs in this front rank⁴.

As pointed out by the paper's authors¹, the abiding mystery of *A. mellifera* is how creatures as tiny as worker bees, with brains containing only a millionth the number of neurons as do ours, are able to perform so many tasks and integrate them into a harmonious whole. Since the first publication on honeybees of the modern era, Charles Butler's *Feminine Monarchie* (1609), discoveries have flowed at the organism and colony levels, justifying

Karl von Frisch's remark that, to scientists, "The life of bees is like a magic well. The more you draw from it, the more there is to draw."

The most celebrated characteristic of *A. mellifera*, next to its honey and pollination services, is, of course, the waggle dance. Foraging workers, on returning to the hive after successful searches for food sources or new nest sites, run figures-of-eight on the vertical comb surfaces, with the middle segment of their body symbolically representing the flight to be taken outward. This latter 'waggle run' contains information about the direction of the target with reference to the Sun, as well as conveying distance from the hive. Timed buzzing and odour secretion enhance the message. Circular 'round dances' supplant the full waggle dance to inform nestmates that the target is close to the nest.

Research in recent years has revealed other performances on the honeybee dance card. If the returning foragers discover many food handlers unemployed as they unload their harvest, they perform the 'shaking dance' to bring more workers on to the dance floor and thence out to the field. If the reverse occurs, in other words if the foragers have trouble passing on their load, they engage in 'tremble dances', to draft in more bees to act as food handlers.

In addition to their terpsichorean programme,

honeybees employ pheromones. These substances, secreted from glands distributed over the body, variously alarm or recruit nestmates, distinguish them from alien bees, and classify them according to gender, caste and age^{4,5}. As workers pass through their natural adult lifespan of about 40 days, their socially active glands variously grow and shrink in programmed time sequences in concert with the labour roles they assume^{6,7}. The progression can be speeded up or reversed according to the needs of the colony⁸. And as they shift among specialities, their receptiveness to particular signals rises or falls.

Finally, worker bees possess an extraordinary memory capacity. They learn the odour of the colony to which they belong. On foraging flights, they use landmarks as well as the instructions given by their dancing nestmates. They can recall the location of up to five flower-beds or other food sites, together with the approximate time of day in which each is most productive^{8,9}.

The decoders of the honeybee genome have begun to address these remarkable social traits at the molecular-genetic level (Box 1, overleaf). From work by generations of previous researchers, they were already aware that virtually every biological system has been

altered to some degree. A first reading of the genome reveals, not unexpectedly, that some of the genes have been modified from ancient precursors. For example, a cluster descended from a single progenitor gene that encoded a member of the yellow protein family here prescribes the royal jelly used in caste determination and queen production. Others, including those that programme chemoreception and food management, seem to include innovations that have evolved since the bee lineage split from that of other insects.

This DNA sequence is a major step towards answering a basic question of social evolution: at the genomic level, what does it take to engineer an advanced colonial insect? As this work is extended, it will soon come to address a second, equally important question: what does it take to make a eusocial insect species in the first place? (In eusocial colonies, members comprise overlapping generations that are divided into reproductive and worker castes to care for the young.) Fortunately, a large fund of information can be gathered to settle this issue, because among the 16,000 or so living bee species known, some are solitary but close to the threshold of eusociality, others have barely made it across the threshold, a few have reverted back to the solitary state, and still others have settled at various degrees of social organization intermediate to the honeybee grade. The evolutionary histories of most of these lines have been worked out on the basis of anatomy¹⁰, and in a few cases with the aid of partial molecular evidence. Among those closest to the honeybees are the solitary euglossines, the bumblebees and the meliponine stingless bees, the last of which seem comparable in complexity of social organization to the honeybees.

Box 1 | The honeybee genome and social lifestyle

The honeybee is the third insect to have its genome sequenced, following the fruitfly (*Drosophila*) and mosquito (*Anopheles*). We can therefore compare the honeybee gene number and content with those of the other two insects, which don't form colonies, to make some guesses at how its social lifestyle has evolved from and in turn shaped its genome.

Honeybees have more genes involved in producing royal jelly, which makes sense as they make it to feed their young, whereas the other two insects don't. They also have far more genes encoding odorant receptors — mirroring the importance of pheromones in sensory communication during the

various bee dances, as well as in distinguishing different castes and bees alien to the colony.

A communal lifestyle lessens some of the hazards faced by a solitary insect. For instance, a hive environment means the honeybee can get away with a simpler outer cuticle than the other insects, and so it has fewer genes encoding cuticle proteins. It also has fewer taste receptors, possibly because most bees feed where another bee has already eaten, reducing the likelihood of ingesting poisonous food.

But not all the differences between the genomes are so easily explained. It seems odd that honeybees have fewer genes involved in immunity

than the other insects. These bees live in crowded quarters, and diseases are easily transmitted in a small space. Perhaps novel pathways help them resist disease, or maybe they are protected sufficiently by social behaviours such as grooming.

The genome paper also reports initial studies of changes in gene expression during the honeybee lifespan. Significantly, developmental transitions, such as that at 2–3 weeks when bees begin to work outside the hive, are accompanied by altered activity of genes expressed in the brain. Moreover, there are prominent differences between the castes in terms of which metabolic genes are expressed.

Chris Gunter

As pieces of this great mosaic are put in place with the aid of comparative genomics, a remarkable history will emerge, yielding many perspectives in developmental evolution and sociobiology.

Edward O. Wilson is at the Museum of Comparative Zoology, Harvard University, 26 Oxford Street, Cambridge, Massachusetts 02138-2902, USA.
e-mail: ewilson@oeb.harvard.edu

1. The Honeybee Genome Sequencing Consortium *Nature* **443**, 931–949 (2006).
2. Wilson, E. O. *Sociobiology: The New Synthesis* (Harvard Univ. Press, 1975).

3. Choe, J. C. & Crespi, B. J. (eds) *The Evolution of Social Behavior in Insects and Arachnids* (Cambridge Univ. Press, 1997).
4. Seeley, T. D. *The Wisdom of the Hive* (Harvard Univ. Press, 1995).
5. Michener, C. D. *The Social Behavior of the Bees: A Comparative Study* (Harvard Univ. Press, 1974).
6. Sekiguchi, F. & Sakagami, S. *Rep. Hokkaido Natl Agric. Exp. Sta.* **69**, 1–65 (1966).
7. Lindauer, M. *Communication Among Social Bees* (Harvard Univ. Press, 1961).
8. Röscher, G. A. Z. *Vergl. Physiol.* **12**, 1–71 (1930).
9. von Frisch, K. *The Dance Language and Orientation of Bees* (Harvard Univ. Press, 1967).
10. Michener, C. D. *The Bees of the World* (Johns Hopkins Univ. Press, 2000).

PALAEOCEANOGRAPHY

In hot water

Christina L. De La Rocha

There has long been scepticism about the geochemical evidence that the ancient ocean was markedly warm. A fresh approach bolsters the case for an ocean that, in the distant past, was indeed quite hot.

Reconstructing past ocean temperatures is an obsession for many Earth scientists. That's understandable — such data provide insights into climate and its links to biogeochemical cycling, ocean circulation and tectonics, not to mention mass extinctions, evolutionary radiations, and other ups and downs of the biosphere. It is the isotopic and trace-element compositions of sedimentary materials that provide the necessary information to quantitatively infer past temperatures. But that task is not straightforward, even for the past few hundred million years for which we have

relatively pristine sediments to analyse. Go back to the Precambrian — roughly 4.5 billion to 0.5 billion years ago — and the job is even more daunting. Not only are there few sedimentary materials of that age left to work with, but all of those that do remain have been altered since they were laid down.

Thus, for decades we have not known what to make of geochemical estimates^{1–3} indicating that, between 3.5 billion and 1.2 billion years ago, the ocean was as hot as 55–85 °C. Those estimates have been especially difficult to believe given the geological evidence for

glaciations during this time⁴. Robert and Chaussidon⁵ (page 969 of this issue), however, have now used the silicon-isotopic composition of ancient rock samples to support the earlier, intriguing results^{1–3}.

Deducing the environmental conditions of the Precambrian ocean would be a stunning achievement. This is, after all, the setting for the origin of life and the evolution of the great microbial diversity that led to and still supports plant and animal life today. The previous estimates of Precambrian ocean temperatures^{1–3} are based on the oxygen-isotopic composition of chert, a siliceous rock that is formed when silica precipitates from sea water. Precipitated silica will have an ¹⁸O to ¹⁶O ratio that is higher than the fluid from which it formed, and the extent to which the ¹⁸O to ¹⁶O ratio is higher is inversely related to temperature. For a given fluid composition, then, the higher the temperature, the lower the resulting ¹⁸O to ¹⁶O ratio of the silica (reported as $\delta^{18}\text{O}$ when normalized to a standard). Thus, the remarkably low $\delta^{18}\text{O}$ values of Precambrian chert, and their

increase towards the Precambrian–Cambrian boundary, could indicate an ancient ocean that was scores of degrees warmer than the modern one and that cooled gradually. But caution dictates ascribing the strikingly low $\delta^{18}\text{O}$ values of ancient chert to other factors, such as the hydrothermal, metamorphic or other alteration of those rocks over billions of years, and not to superwarm Precambrian sea water.

But the low $\delta^{18}\text{O}$ values may indicate high temperatures after all. Robert and Chaussidon⁵ point out that the Precambrian chert samples with the highest $\delta^{18}\text{O}$ values (that is, those least likely to have been altered) have $\delta^{18}\text{O}$ values that correlate strongly with their silicon-isotopic composition ($^{30}\text{Si}/^{28}\text{Si}$, reported as $\delta^{30}\text{Si}$). It is unlikely that such a relationship would have been maintained during alteration, suggesting that the isotopic signal of these particular chert samples (Fig. 1) is primary and robust. This evidence alone supports the contention that the Precambrian ocean was quite warm. But Robert and Chaussidon⁵ have taken things a step further by using the silicon isotopes on their own to make an alternative temperature estimate.

The $\delta^{30}\text{Si}$ values of chert older than 1 billion years can be high^{5,6} relative to the siliceous sediments and ocean waters of the modern day^{7,8}. Precambrian concentrations of dissolved silicon were extreme, not far from the point of saturation, and were controlled largely by abiotic reactions⁹. Robert and Chaussidon assume that seawater $\delta^{30}\text{Si}$ was steady over time, and argue that the $\delta^{30}\text{Si}$ of chert and that of the silica precipitated during the circulation of hydrothermal fluids through ocean crust must have together equalled the $\delta^{30}\text{Si}$ of the mantle, the ultimate source of silicon to the system. Because precipitates have a lower $\delta^{30}\text{Si}$ than the dissolved silicon from which they formed, an increase in the proportion of hydrothermal silicon precipitated should be reflected in an increase in the $\delta^{30}\text{Si}$ of chert. And if the temperature of sea water controlled the extent of the hydrothermal silicification, the $\delta^{30}\text{Si}$ of Precambrian chert should reveal the temperature of the ocean at that time.

As with any model built to interpret isotopic results in the context of (bio)geochemical cycling, there are numerous assumptions at work here. One is that dissolved silicon input to the ocean has the same $\delta^{30}\text{Si}$ signal as that of Earth's mantle. This is not true in modern times, where the $\delta^{30}\text{Si}$ of the silicon input reflects the sequestration of silicon into clay minerals during weathering on the continents^{8,10}. Ignoring clay formation might be fine for the part of the Precambrian lacking sufficient continental mass for the large-scale formation of clays. But this is not true for the entire Precambrian. Continental weathering and clay formation undoubtedly influenced the $\delta^{30}\text{Si}$ values of the younger Precambrian chert, throwing the modelled temperature estimates out of kilter.

Nonetheless, Robert and Chaussidon's



Figure 1 | Memory store. This example of a chert analysed by Robert and Chaussidon⁵ comes from the Jinxian group, China, and is some 1.4 billion years old. Microcrystalline quartz within a sample such as this may have retained an isotopic memory of the conditions under which the chert initially formed — evidence that can be unlocked using the fine sampling beam of the ion microprobe.

M. CHAUSSIDON

temperature estimates⁵ using silicon isotopes are in remarkable agreement with the earlier estimates based on oxygen isotopes^{1–3}. Both records suggest that average ocean temperatures were between 60 °C and 80 °C for much of the time between 3.5 billion and 2.0 billion years ago. And both records imply cooling over the Precambrian, with more conventionally hospitable ocean temperatures of 20–30 °C reached by 1.5 billion years ago.

A Precambrian ocean that was at times warmer than 70 °C certainly gives pause for thought. No photosynthetic organism that can grow above 73 °C has ever been found¹¹. Moreover, the declining solubility of oxygen gas with temperature would have exacerbated ocean hypoxia under the low oxygen levels of the Precambrian atmosphere, potentially impeding the development of multicellular animal life³. The silicon-isotope data from the Precambrian are striking. But given these points, and the existence of glacial deposits within the Precambrian, it will probably take

more than $\delta^{18}\text{O}$ and $\delta^{30}\text{Si}$ together to convince sceptics of a hot-tub Precambrian sea. ■

Christina L. De La Rocha is at the Alfred Wegener Institute for Polar and Marine Research, Columbusstraße, 27568 Bremerhaven, Germany. e-mail: crocha@awi-bremerhaven.de

1. Knauth, L. P. & Lowe, D. R. *Earth Planet. Sci. Lett.* **41**, 209–222 (1978).
2. Knauth, L. P. & Lowe, D. R. *Geol. Soc. Am. Bull.* **115**, 566–580 (2003).
3. Knauth, L. P. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **219**, 53–69 (2005).
4. Kopp, R. E., Kirschvink, J. L., Hilburn, I. A. & Nash, C. Z. *Proc. Natl Acad. Sci. USA* **102**, 11131–11136 (2005).
5. Robert, F. & Chaussidon, M. *Nature* **443**, 969–972 (2006).
6. Ding, T. et al. *Silicon Isotope Geochemistry* (Geol. Publ. House, Beijing, 1996).
7. Douthitt, C. B. *Geochim. Cosmochim. Acta* **46**, 1449–1458 (1982).
8. De La Rocha, C. L., Brzezinski, M. A. & DeNiro, M. J. *Geochim. Cosmochim. Acta* **64**, 2467–2477 (2000).
9. Siever, R. in *Scientists on Gaia* (eds Schneider, S. H. & Boston, P. J.) 287–295 (MIT Press, 1991).
10. Ziegler, K., Chadwick, O. A., Brzezinski, M. A. & Kelly, E. F. *Geochim. Cosmochim. Acta* **69**, 4597–4610 (2005).
11. Brock, T. D. *Science* **230**, 132–138 (1985).

PALAEONTOLOGY

Modern look for ancient lamprey

Philippe Janvier

It was once thought that lampreys evolved from armoured jawless vertebrates. But a recently discovered lamprey fossil dates from the twilight age of their supposed ancestors, and looks surprisingly modern.

Most living vertebrates have jaws. Such creatures comprise about 51,000 species ranging from sharks to four-legged land vertebrates. In contrast, living jawless vertebrates — known as cyclostomes — consist of only 60 species. These fall into two groups: hagfishes and lampreys (Fig. 1, overleaf). Both jawless groups lack paired fins, and many of their anatomical and physiological features are regarded as primitive relative to those of their jawed counterparts. As cyclostomes also lack scales and bone, their history is mostly undocumented by fossils, but a number of recent findings

have started to fill in the gaps. On page 981 of this issue, Gess *et al.*¹ report the oldest fossil lamprey to date, a discovery that calls for a reassessment of cyclostome evolution.

Until the late twentieth century, cyclostomes were assumed to be 'degenerate' descendants of armoured jawless vertebrates — known informally as ostracoderms — that lived from the Ordovician period to the Devonian period, 490 million to 358 million years (Myr) ago (Fig. 1). Hagfishes and lampreys were thought to share a common ancestor that derived from certain ostracoderms — either osteostracans

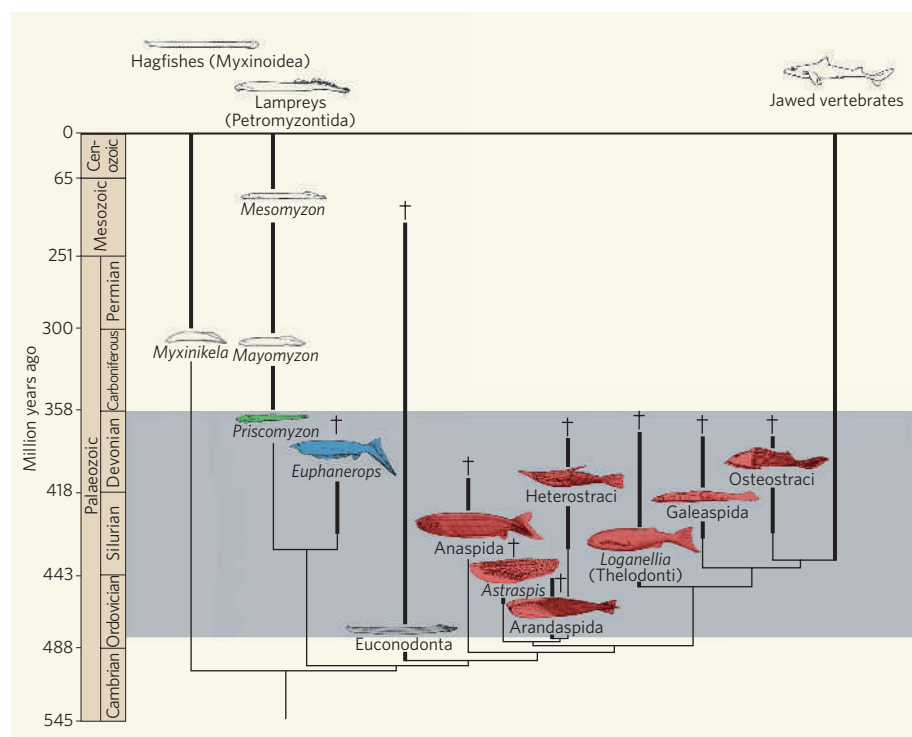


Figure 1 | Vertebrate tree. Lampreys and hagfishes are the only two living groups of jawless vertebrate. The 360-million-year-old lamprey *Priscoomyzon* (green) discovered by Gess *et al.*¹ is very similar to modern lampreys, even though it dates from the twilight age (grey area) of the armoured jawless vertebrates (known as ostracoderms, in red) that were once considered to be ancestors of hagfishes and lampreys^{2,15}. The evolutionary tree now proposed by Gess *et al.* (simplified version shown here; crosses indicate extinction dates) agrees with the current consensus that ostracoderms are more closely related to jawed vertebrates than to lampreys or hagfishes^{1,7-9,12,14}. This suggests that living jawless vertebrates and their forerunners never developed an extensive bony skeleton¹⁴, and that their origin must lie among early Palaeozoic jawless vertebrates that lacked scales and bone, such as *Euphanerops*^{1,13} (blue).

or anaspids² (Fig. 1). The evolutionary divergence of the two cyclostome groups from each other was regarded as relatively recent, possibly occurring during the Mesozoic period 251–65 Myr ago. So the subsequent discovery of 300–330-Myr-old fossil lampreys and hagfishes from the Carboniferous period came as a surprise³. These fishes are exceptionally preserved as soft-tissue imprints and are almost identical to their modern successors. Nevertheless, there remained a gap of at least 35 Myr between these early cyclostomes and their previously supposed ancestors from the Devonian period, during which time they could have evolved by losing their mineralized skeleton.

But Gess and colleagues' fossil lamprey¹, *Priscoomyzon*, actually dates from the late Devonian period. *Priscoomyzon* shows characteristic lamprey features, such as a grotesquely large sucker armed with horny teeth that surrounds the mouth, and a basket-like gill skeleton. This shows that lamprey morphology has been astonishingly stable for 360 Myr, and proves that lampreys and hagfishes had already diverged by late Devonian times, earlier than previously thought. Admittedly, *Priscoomyzon*, like the Carboniferous lampreys, differs from its modern equivalents in minor details; compare this with the recently described 125-Myr-old *Mesomyzon*⁴, the first Mesozoic lamprey, which would probably go unnoticed in a

present-day brook (Fig. 1). The large sucker of *Priscoomyzon* suggests that, like some modern-day lampreys, it could fasten on to other fishes and suck their blood. Yet only 19 living lamprey species (out of 38) feed this way⁵. Other lampreys mainly use their sucker to either secure themselves while at rest or carry stones for nest building.

The relationships between living hagfishes, lampreys and jawed vertebrates are hotly debated, because of conflicting distributions of morphological and physiological traits on the one hand, and of DNA and RNA sequence data on the other. The morphological and physiological aspects suggest that lampreys (but not hagfishes) are the sister group of jawed vertebrates^{1,3,6-9}, whereas gene sequences generally suggest that lampreys and hagfishes are sister groups^{10,11}. Fossils sometimes help to resolve such conflicts, by revealing combinations of traits in an extinct species that better support a particular relationship. Frustratingly, *Priscoomyzon* does not help in resolving the problem of lamprey relationships, because it provides no new informative combinations of characteristics compared with post-Devonian and extant lampreys.

Morphology-based evolutionary trees of living and fossil vertebrates have long been prone to change. The tree diagrams yielded by computer programs since the early 1990s



50 YEARS AGO

Graduate Employment: A Sample Survey — The desirability of information of this kind has been recognized ever since 1946, when the Barlow Committee's Report on Scientific Man-power sowed the seeds for what has become a widespread concern about Britain's chronic shortage of scientists and engineers; about the deleterious effect of the shortage on our national position; about the scarcity of teachers of science and mathematics; and about the failure of our educational system to weave into the pattern of general culture an appreciation of scientific matters and more important, an awareness of the incredible speed with which their influence over human affairs is growing.

From *Nature* 27 October 1956.

100 YEARS AGO

"The recent radium controversy" — I was absent from Montreal during the time of the interesting discussion which appeared in the *Times*... In the course of this discussion some weight has been attached to a remark in the second edition of my book "Radio-activity" viz. that radium is a compound of helium and lead... Lord Kelvin quite correctly quotes my words, but I feel that the statement is liable to leave an erroneous impression of my views on the question... At the risk of being somewhat lengthy, I should like to quote fully some statements made in my book... "If the α particle is a helium atom, at least three α particles must be expelled from uranium (238.5) to reduce its atomic weight to that of radium (225). It is known that five α particles are expelled from radium during its successive transformations. This would make the atomic weight of the final residue $225 - 20 = 205$. This is very nearly the atomic weight of lead, 206.5. I have for some time considered it probable that lead is the end or final product of radium." ... I think that the above quotation makes my position clear on this subject. **E. Rutherford**

From *Nature* 25 October 1906.

50 & 100 YEARS AGO

generally show hagfishes and lampreys diverging from other vertebrates before the ostracoderms (Fig. 1). In other words, ostracoderms, although jawless, are more closely related to jawed vertebrates than to cyclostomes^{3,7–9}. This would mean that lampreys and hagfishes diverged away from jawed vertebrates before the advent of the earliest ostracoderms 477 Myr ago, at the latest^{8,9,12}. DNA and RNA data suggest that this divergence occurred 535–462 Myr ago, assuming that lampreys and hagfishes are sister groups to each other¹¹.

So, it is not too surprising that lampreys turn up in the Devonian period, 360 Myr ago. What is surprising is that they are already very similar to modern lampreys. What, then, did earlier or more primitive lampreys look like? The tree proposed by Gess *et al.*¹ (Fig. 1) shows the late Devonian (370 Myr) *Euphanerops* as a sister group of lampreys. This 'naked' jawless vertebrate somewhat resembles lampreys with its basket-like gill apparatus, gill pouches and possibly similar cartilage structure¹³. But it also resembles the scale-covered anaspids (Fig. 1), which form one of the ostracoderm groups^{2,3}. In current vertebrate evolutionary trees, *Euphanerops* and anaspids fall alongside the other ostracoderms^{8,9}, but in Gess and colleagues' tree¹, only anaspids appear as one of the earliest groups on this stem of the jawed-vertebrate branch, preceding the other ostracoderms, while *Euphanerops* appears as a lamprey relative (Fig. 1). *Euphanerops* is likely to be a late representative of an older group whose anaspid-like aspect may reflect that of the common ancestor of lampreys and jawed vertebrates. Primitive lampreys possibly resembled such anaspid-like forms, but devoid of scales and bone.

It seems that an early segment of vertebrate evolution, before the rise of bone, is documented only by fishes that lacked a mineralized skeleton (except for rare occurrences of slightly calcified cartilage). For this reason, exceptional discoveries of Palaeozoic soft-bodied fishes such as *Priscomyzon* are of tremendous value¹⁴.

Philippe Janvier is at the Muséum National d'Histoire Naturelle, UMR 5143, CNRS, 8 Rue Buffon, Paris 75005, France.
e-mail: janvier@mnhn.fr

1. Gess, R. W., Coates, M. I. & Rubidge, B. S. *Nature* **443**, 981–984 (2006).
2. Heintz, A. in *The Biology of Myxine* (eds Brodal, A. & Fänge, R.) 9–31 (Universitetsforlaget, Oslo, 1963).
3. Janvier, P. *Early Vertebrates* (Oxford Univ. Press, 1996).
4. Chang, M.-M., Zhang, J. & Miao, D. *Nature* **441**, 972–974 (2006).
5. Hardisty, M. W. *Lampreys: Life without Jaws* (Forrest Text, Ceredigion, 2006).
6. Løvtrup, S. *The Phylogeny of the Vertebrata* (Wiley, New York, 1977).
7. Gagnier, P. Y. *Ann. Paléont. (Vert.)* **79**, 119–166 (1993).
8. Donoghue, P. C. J., Forey, P. L. & Aldridge, R. J. *Biol. Rev.* **75**, 191–351 (2000).
9. Donoghue, P. C. J. & Smith, M. P. *Trans. R. Soc. Edinb.* **92**, 15–37 (2001).
10. Furlong, R. F. & Holland, P. W. H. *Zool. Sci.* **19**, 593–599 (2002).
11. Hedges, S. B. in *Major Events in Early Vertebrate Evolution*

(ed. Ahlberg, P. E.) 119–134 (Taylor & Francis, London, 2001).

12. Donoghue, P. C. J., Smith, M. P. & Sansom, I. J. in *Telling Evolutionary Time: Molecular Clocks and the Fossil Record* (eds Donoghue, P. C. J. & Smith, M. P.) 191–223 (Taylor & Francis, London, 2003).

13. Janvier, P., Desbiens, S., Willett, J. A. & Arsenault, M. *Nature* **440**, 1183–1185 (2006).
14. Donoghue, P. C. J. & Sansom, I. J. *Microsc. Res. Tech.* **59**, 352–372 (2002).
15. Jarvik, E. *Basic Structure and Evolution of Vertebrates* (Academic, London, 1980).

ELECTROMAGNETISM

Like a speeding watch

Miles Padgett

A rotation in light's electric-field vector can alter the light's frequency. This rotational equivalent of the Doppler effect has proved surprisingly elusive, but has now been spotted in the laboratory.

When an observer moves towards or away from a light source, the measured frequency of the light shifts in proportion to the source and observer's relative velocity. Perhaps the best-known example of this Doppler shift is the 'redshift' — a displacement towards lower frequencies — of light emitted by receding galaxies. The Doppler shift also applies to sound waves, causing the characteristic rise and fall in the pitch of a passing police siren. As they report in *Physical Review Letters*¹, Barreiro and colleagues have now measured the broadening of a spectral line caused by an effect analogous to the conventional Doppler effect. In this case, however, the shift in frequency arises not from linear motion, but from rotation — the rarely encountered rotational Doppler effect.

Place a watch at the centre of a rotating turntable, and, viewed from above, its hands will seem to rotate more quickly (Fig. 1). This classical effect applies to all rotating vectors, for example to the spatial pattern of the electric field of any light beam carrying angular momentum. The electric-field vector rotates at the frequency of the light, but an additional rotation of the beam about its axis of propagation will speed up or slow down the field rotation, resulting in a frequency shift proportional to the rate of rotation of the beam.

Importantly, the rotational Doppler effect is seen by looking at a rotating body along or parallel to its axis of rotation. It is therefore distinct from the linear Doppler shift that arises from viewing the edges of an extended body — a galaxy, say — in a direction perpendicular to its rotation axis. The effect was originally observed and analysed in terms of Jones polarization matrices, which describe the effect of a medium on the orientation of the electric-field vector². The effect has also been cited as an example of a geometric, or Berry, phase shift that occurs in a system whose parameters are progressively changed before it is brought back to its initial state³.

At a more subtle level, the angular momentum of light can be divided into spin and orbital components⁴. The spin angular momentum is associated with the rotation of the electric-field



Figure 1 | Is that the time? Viewed from above, the hands of a spinning watch will seem to rotate more quickly than normal — the rotational Doppler effect investigated by Barreiro *et al.*¹.

vector. This corresponds to circular polarization, meaning that the tip of the vector traces a circular pattern in space as it rotates. The orbital angular momentum, on the other hand, is associated with a rotation of the light wave's phase. Whereas the spin angular momentum can take only two independent states (clockwise or anticlockwise, according to the sense of the field vector's rotation), the orbital angular momentum can take any number of states with different values of angular momentum. In all cases, for a given field rotation speed, the Doppler frequency shift is proportional to the total angular momentum. This shift has been directly quantified for the rotation of electromagnetic beams at millimetre wavelengths⁵.

Barreiro *et al.*¹ examined the spectrum of transitions between energy levels in rubidium gas, using as probes two light beams carrying different values of orbital angular momentum. The rotational Doppler effect is easily masked by larger frequency shifts, such as the normal linear Doppler shift. The authors used a geometry in which all these unwanted frequency shifts balanced to zero, and chose an intricate rubidium transition that needs two photons and a magnetic field to be activated. The transition had an initial linewidth of

52 kilohertz, but when orbital angular momentum was introduced to the light beams, this broadened to 300 kHz — a clear signal of the rotational Doppler effect.

Although such phenomena are based purely on the equations of nineteenth-century physics, it has only been realized in the past two decades that a light beam generated in the laboratory can carry a distinct orbital angular momentum, in addition to its spin. Many resulting phenomena have been reported⁶ in both the classical and quantum regimes: optical spanners, for instance, in which light's angular momentum causes microscopic objects to both

spin and orbit; nonlinear effects whereby the conservation of orbital angular momentum causes the light beam's phase to change; and Heisenberg-type uncertainty relationships, where measurement of angular position sets the orbital angular momentum of light.

The rotational Doppler effect is particularly interesting because here the spin and orbital components of the angular momentum are indistinguishable; instead it is the total angular momentum of the light beam that is crucial. Astoundingly, more than 100 years after its formulation, classical electromagnetism still has surprises in store. ■

Miles Padgett is in the Department of Physics and Astronomy, Kelvin Building, University of Glasgow, Glasgow G12 8QQ, UK.
e-mail: m.padgett@physics.gla.ac.uk

1. Barreiro, S., Tabosa, J. W. R., Failache, H. & Lezama, A. *Phys. Rev. Lett.* **97**, 113601 (2006).
2. Garetz, B. A. & Arnold, S. *Opt. Commun.* **31**, 1 (1979).
3. Simon, R., Kimble, H. J. & Sundarshan, E. C. G. *Phys. Rev. Lett.* **61**, 19–22 (1988).
4. Allen, L., Beijersbergen, M. W., Spreeuw, R. J. C. & Woerdman, J. P. *Phys. Rev. A* **45**, 8185–8189 (1992).
5. Courtial, J., Robertson, D. A., Dholakia, K., Allen, L. & Padgett, M. J. *Phys. Rev. Lett.* **81**, 4828–4830 (1998).
6. Padgett, M., Courtial, J. & Allen, L. *Phys. Today* **57**, 35–40 (2004).

GENOMICS

Blueprints for partnerships

David A. Stahl and Seana K. Davidson

Gutless marine worms harness the resources of a team of bacteria in lieu of a digestive or excretory system. A genome-sequence analysis now defines the roles of the microbes.

Earth was a planet of microbes long before multicellular life appeared and diversified. The evolution of all plants, animals and fungi has thus been profoundly influenced by micro-organisms. As a result, many innovative partnerships, known as symbioses, have developed between microbes and their host species^{1,2}. On page 950, Dubilier and colleagues (Woyke *et al.*)³ use the power of genome sequencing to reveal a multifaceted partnership between a worm and four types of bacterium*.

The most closely studied relationships between animals and microbes are those that cause diseases in animals. But there are many beneficial associations that are much less studied and appreciated. The human body contains ten times more microbial cells than human cells, and our coexistence with microbes is generally thought to be useful — perhaps even essential — for our development and health^{4,5}. Such beneficial interactions can have many functions. In all cases of long-established associations, both the host and the microbe coevolve. The resulting evolutionary changes can be so extensive that the association becomes essential to both host and microbe 'symbiont'. It is especially challenging to understand the complex exchanges between the partners and to work out which organism does what, as the partners cannot survive separately.

Small, gutless marine worms, distantly related to earthworms, are involved in a remarkable example of a highly coevolved symbiosis. Found in coastal sediments, these worms lack a mouth, digestive tract or typical excretory system^{6,7}. Instead, they harbour bacteria immediately below their outer skin (Fig. 1). These bacteria

can oxidize or reduce sulphur compounds to generate the energy used to convert carbon dioxide into organic carbon, which the worm then feeds on for growth⁸.

The association between the worm and its symbionts is more complex than initially thought. By sequencing the bacterial ribosomal RNA genes, Dubilier and colleagues^{9,10} previously discovered that the worms stably support at least five distinct species of bacteria. Bacteria that obtain energy by oxidizing sulphides coexist with those that grow in the absence of oxygen. The latter obtain energy

by 'breathing' sulphate, producing sulphide as a result of this respiration rather than water. Other bacteria known as spirochaetes are also present, although their function is enigmatic. The authors' model of how these bacteria work together to feed the worm was initially based on measurements of the intact worm–bacteria system and the sequences of a few functional genes, but crucial details were lacking. Dubilier and colleagues³ now infer the metabolic interactions sustaining this remarkable collaboration by using a genome-sequence analysis.

The authors used DNA obtained from the bacteria in the worm tissue to construct a library of DNA fragments. This library was a random mixture of fragments obtained from all five symbionts. The DNA fragments were then examined to determine the sequences of their nucleotide building-blocks. Fragments collectively containing about 204 million nucleotides were analysed in this way. Sequences of about 3,000 nucleotides each were then assigned to different genome types using a method known as nucleotide-frequency binning. In this way,

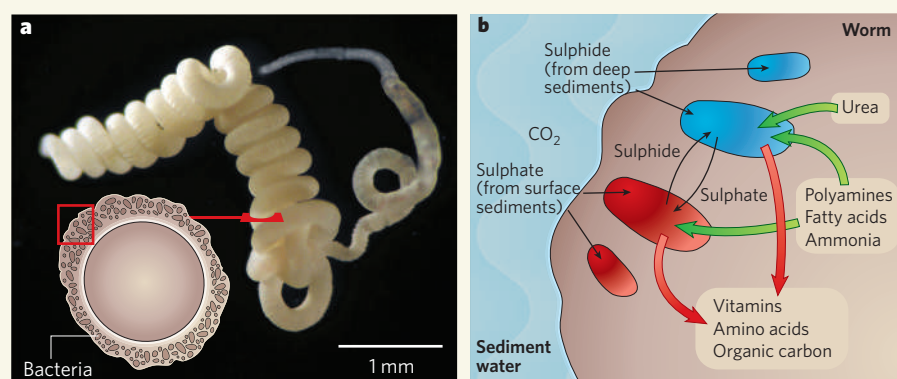


Figure 1 | A bacterial organ. **a**, Gutless marine worms that live in coastal sediments lack digestive and excretory systems. Instead, they rely on bacteria beneath their outer skin, as shown in the cross-section of the worm. **b**, Dubilier and colleagues³ determined the combined genome sequences of four of these bacterial species to identify the potential metabolic processes and exchanges occurring in the worm. The four species come in two varieties: those that oxidize sulphides (blue) to generate energy, and those that reduce sulphates (red) for the same purpose. The energy is used to convert carbon dioxide from the environment into organic carbon on which the worm feeds, along with vitamins and amino acids (red arrows). Sulphides and sulphates may be obtained externally from different sediments, or as metabolic products from the other bacteria (black arrows). The bacteria in turn receive compounds from the worm (green arrows). These include molecules that are useful to the bacteria, such as fatty acids and polyamines. The bacteria also absorb the toxic waste products urea and ammonia, so the worm does not need to excrete these compounds. (Photo courtesy of C. Lott.)

*This article and the paper concerned³ were published online on 17 September 2006.

the authors assembled two nearly complete genomes and two partial genomes of the predominant symbionts. Gene function in the symbionts was inferred by identifying closely related genes in other well-characterized bacteria. With this information in hand, Dubilier and colleagues began to unravel the metabolic interactions that form this multi-species symbiosis.

The four characterized bacterial species seem to form a versatile system for supplying the host with organic carbon, amino acids and vitamins, and for eliminating nitrogenous waste. The authors identified bacterial genes that encode proteins that take up the nitrogen-containing compounds urea and ammonia, thus allowing the worm to conserve nitrogen rather than excreting these toxic waste products, and so possibly explaining the lack of excretory organs in the host.

Co-dependent interactions between the sulphide-oxidizing symbionts seem to be complex. All four symbionts can convert CO₂ into organic carbon, a process for which three distinct pathways exist, and the sulphate-reducing bacteria can also feed the host by taking up dissolved organic compounds from the environment. Together, the symbionts provide the worm with the enzymatic machinery to exploit various organic and inorganic energy sources, and to respire using molecules other than oxygen, such as sulphates, nitrates and fumarate.

The authors propose³ that sulphide produced by the sulphate-reducing bacteria acts as an energy source for the sulphide-oxidizing bacteria, so creating an internal sulphur cycle that sustains both symbionts and also nourishes the worm. However, the compounds needed to drive this cycle may not all coexist in the same environment, as the preferred oxidants (oxygen or nitrate) for sulphide are found in surface sediments, whereas sulphate respiration is an anaerobic process that is only possible in deeper sediments that lack oxygen. It seems that the worm assists the bacteria by supplying another oxidant: the genomes show that the sulphide-oxidizing symbionts can respire using trimethylamine *N*-oxide (TMAO), a compound produced by the worm. This allows the bacteria to grow in the absence of oxygen or nitrate.

Furthermore, the host can move between the surface sediments and the deeper regions, so that an understanding of how the bacteria generate energy must also consider the worm's behaviour. One of the sulphide-oxidizing bacteria might partially oxidize sulphide using TMAO and store the resulting sulphur in globules for later use. When the worm migrates to the surface, this stored sulphur could then be fully oxidized to sulphate, and so the cycle continues. The sulphate-reducing bacteria, in turn, possess genes that enable them to grow in deep, oxygen-starved sediments by using molecular hydrogen to reduce sulphate and provide the host with carbon.

But the story remains incomplete. It is

unclear why several species of sulphur-cycling bacteria are required, or what features of their co-dependent interactions sustain them. It also remains to be seen how the host fits into this web of interactions — for example, can the symbionts modify the worm's behaviour to direct it to different depths of the sediments? And what is the role of the fifth microbial partner in this complex symbiosis, the spirochaete bacterium? Studies of these 'lowly' worms should improve our understanding of many complex symbioses, including those found in the human body. ■

David A. Stahl and Seana K. Davidson are in the Department of Civil and Environmental Engineering, University of Washington, More Hall, Box 352700, Seattle,

Washington 98195-2700, USA.

e-mails: dastahl@u.washington.edu;
skdavid@u.washington.edu

1. Margulis, L. & Fester, R. (eds) *Symbiosis as a Source of Evolutionary Innovation: Speciation and Morphogenesis* (MIT Press, Cambridge, MA, 1991).
2. Seckbach, J. (ed.) *Symbiosis: Mechanisms and Model Systems* (Kluwer Academic, Norwell, MA, 2002).
3. Woyke, T. *et al. Nature* **443**, 950–955 (2006).
4. Hooper, L. V., Midtvedt, T. & Gordon, J. I. *Annu. Rev. Nutr.* **22**, 283–307 (2002).
5. Hooper, L. V. *Trends Microbiol.* **12**, 129–134 (2004).
6. Giere, O. *Cah. Biol. Mar.* **20**, 301–314 (1979).
7. Giere, O. *Mar. Ecol. Prog. Ser.* **5**, 353–357 (1981).
8. Dubilier, N. *et al. Nature* **411**, 298–302 (2001).
9. Dubilier, N., Blazejak, A. & Ruhlan, C. *Prog. Mol. Subcell. Biol.* **41**, 251–275 (2006).
10. Blazejak, A., Erseus, C., Amann, R. & Dubilier, N. *Appl. Environ. Microbiol.* **71**, 1553–1561 (2005).

GEOPHYSICS

Protons lead the charge

Greg Hirth

Is the anomalously high electrical conductivity seen in part of Earth's mantle caused by protons derived from hydrous defects in the mineral olivine? Two groups investigate this possibility — and draw different conclusions.

Many lines of evidence indicate that Earth's dynamic processes are controlled by water — more specifically, by the hydrogen ions (protons) present in minerals in its interior. What has been lacking is a reliable way to determine the concentration and distribution of this hydrogen. Theoretical considerations indicate that the magnitude of a mineral's electrical conductivity puts a constraint on its hydrogen content. In this issue, studies by Yoshino *et al.* (page 973)¹ and Wang *et al.* (page 977)² demonstrate that small amounts of hydrogen increase the electrical conductivity of olivine — a silicate that is the most abundant mineral in Earth's upper mantle — by between 100 and 1,000 times. These experimental results provide an initial verification of a process that has been proposed to explain high electrical conductivity in the upper mantle — but also supply several caveats.

Geoscientists have speculated that the apparently unique features of plate tectonics on Earth — the phenomenon does not seem to occur on any other Solar System body — are controlled by the presence of water. This hypothesis stems from the existence of oceans, the identification of tectonic processes that cycle volatile substances into and out of the mantle, and an appreciation of the influence of water on diffusion within the mantle, and on mantle viscosity and melting. For example, the melting temperature of mantle rocks decreases by some 700 °C in the presence of water³. Furthermore, although the interior of the convecting Earth is too hot for hydrous mineral phases, such as micas, to be stable, small concentrations (around 50 parts

per million by weight) of hydrogen in olivine result in as much as a 30–100-fold reduction in viscosity⁴. At least another ocean's worth of water resides in Earth's interior⁵, and such observations have stimulated studies on exactly how much water there is, and its role in the formation of tectonic plates and in the thermal evolution of Earth 4.5 billion years ago⁶.

A technique known as magnetotellurics is used to study Earth's electrical conductivity by measuring its response to interactions between the solar wind and the ionosphere, the layer of Earth's atmosphere that is ionized by solar radiation. Specifically, electrical conductivity is determined as a function of depth by measuring variations in the magnetic and electric fields at Earth's surface over periods of months to years.

The electrical conductivity measured for some parts of the upper mantle using magnetotellurics is significantly greater than that of the minerals actually present in the mantle. In 1990, Shun-ichiro Karato proposed an explanation⁷ for this apparent discrepancy based on experiments by Mackwell and Kohlstedt on the diffusion of hydrogen ions in olivine⁸. These experiments demonstrated two remarkable features. First, protons are extremely mobile in olivine — almost as mobile as electron holes (a hole is the absence of an electron, and changes ferrous iron to ferric iron, for example). Second, the diffusion rate of protons is exceptionally anisotropic, with the rate parallel to the shortest crystallographic axis (known as the *a* axis) being 50 to 100 times greater than that in other directions.

Karato's insight came from application of the

DEVELOPMENTAL BIOLOGY

Red-eye redirected

In most mammals, the cornea — the transparent part of the eye over the lens — has no blood vessels. This trait is obviously essential for optimal vision, but it also means the cornea is a useful experimental system for studying the factors that promote or inhibit the formation of blood vessels (angiogenesis). But why the cornea remains avascular despite the presence of the potent angiogenic factor VEGF-A and the proximity of other highly vascular tissues has remained unclear.

In this issue, B. K. Ambati *et al.* (*Nature* **443**, 993–997; 2006)

show that in the cornea a soluble receptor, called sflt-1, traps VEGF-A, stopping it from directing the formation of blood vessels. When the authors blocked expression of sflt-1 using a variety of approaches, they saw an increase in free VEGF-A and in corneal vascularization, demonstrating that sflt-1 maintains corneal avascularity.

Blood vessels form spontaneously in the corneas of certain mutant mouse strains, as well as in those of people who suffer from a condition known as aniridia as a result of a similar mutation. Ambati and colleagues show that in all these

cases, corneal vascularization is accompanied by a deficiency in the expression of sflt-1. In the mice, injections of sflt-1 reduced the corneal vascularization.

The authors surveyed different mammalian species to see whether the close relationship between the presence of sflt-1 and an avascular cornea is evolutionarily conserved. Manatees (pictured) are the only known organism with uniformly vascularized corneas. They compensate for their impaired vision with highly developed sensory bristles that enable them to navigate and locate food. Manatees live primarily in turbid freshwater areas, and corneal vascularization may result from or protect against this



K. CALVO/V&W/IMAGEQUESTMARINE.COM

environment. Interestingly, no sflt-1 expression was detected in manatee corneas, whereas those of dugongs (which belong to the order Sirenia, like manatees) and of elephants — the manatee's closest terrestrial relatives — all produced sflt-1.

Barbara Marte

Nernst–Einstein relation, which shows how electrical conductivity changes with the diffusion rate and concentration of charged species. Intriguingly, the *a* axis of olivine (the direction of potentially high conductivity) aligns parallel to the flow during viscous creep in the mantle. Because olivine deforms more easily along some crystallographic axes than along others, the alignment of olivine crystals can be observed using various seismic techniques. In regions of the oceanic mantle for which seismic and geochemical measurements have provided independent constraints on the alignment of olivine and on the water content, the Nernst–Einstein relation provides a good fit to both the magnitude and anisotropy of electrical conductivity. It also explains rapid changes in conductivity that occur in regions where water content is predicted to decrease⁹. Yet despite the apparent success of this approach, the idea that water affects conductivity met with scepticism owing to the lack of experimental verification.

Yoshino *et al.*¹ and Wang *et al.*² both report a huge increase in the conductivity of olivine with increased hydrogen content. But important differences in the results lead the authors to varying conclusions. Yoshino and colleagues' measurements¹, which were conducted on olivine single crystals, show an anisotropy of conductivity similar to that predicted by the hydrogen diffusion measurements. But the measured temperature dependence of conductivity for crystals oriented parallel to the *a* axis is significantly smaller than that of other directions. Thus, at face value the extrapolation of these data to conditions inside Earth (from the maximum experimental temperature of 723 °C to around 1,400 °C) indicates that the hydrogen effect is too small to explain the high conductivity of the mantle. The experiments were conducted at low temperatures to inhibit hydrogen loss from the sample by diffusion.

Wang and colleagues' experiments² were conducted on aggregates of fine-grained olivine

at higher temperatures, and show a somewhat greater effect of hydrogen on conductivity, as well as a slightly greater temperature dependence. Extrapolating their data to mantle conditions provides excellent agreement with the observed conductivity. Based on these initial experiments, the influence of hydrogen concentration on conductivity is lower than that predicted by Karato's original hypothesis⁷, although Wang *et al.* note² that the limited range of hydrogen content in their samples prevents a determination of the precise extent of this difference. Thus, these authors argue that only a fraction of dissolved hydrogen contributes to the conductivity during experiments — a conclusion drawn from similar measurements on wadsleyite, which is a form of olivine found at high pressures¹⁰.

More work is required to resolve the differences between these studies. Variables such as grain size, other mantle mineral phases, aggregate anisotropy and the presence of small amounts of hydrous melt all need to be explored. But it is promising to note how well Karato's application of the Nernst–Einstein relation agrees with Wang and colleagues' data over the temperature range of the hydrogen diffusion experiments (800–1,000 °C; Fig. 2a on page 978). This agreement is even stronger if new constraints on the solubility of hydrogen in olivine are taken into account¹¹: the technique that both Yoshino *et al.* and Wang *et al.* used to measure hydrogen content may underestimate the hydrogen concentration by a factor of three (meaning that the dashed lines in Wang and colleagues' Figure 2a are plotted a factor of three too low).

The discrepancy between the Nernst–Einstein relation and the data at lower temperature might then be explained by a fundamental difference between the conductivity and diffusion studies, as Wang and colleagues note in the supplementary information to their paper. During diffusion experiments,

which measure the rate of hydration of olivine, hydrogen defects must first be created and then migrate. Activation energy must be supplied for both these processes. In the conductivity experiments, by contrast, the samples are 'doped' with hydrogen before the experiment, and activation energy is required only to get the hydrogen moving. As a result, the temperature dependence of the conductivity experiments is lower than that of the diffusion experiments. ■

Greg Hirth is in the Department of Geology & Geophysics, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA. e-mail: girth@whoi.edu

1. Yoshino, T., Matsuzaki, T., Yamashita, S. & Katsura, T. *Nature* **443**, 973–976 (2006).
2. Wang, D., Mookherjee, M., Xu, Y. & Karato, S. *Nature* **443**, 977–980 (2006).
3. Grove, T. L., Chatterjee, N., Parman, S. W. & Médard, E. *Earth Planet. Sci. Lett.* **249**, 74–89 (2006).
4. Mei, S. & Kohlstedt, D. L. *J. Geophys. Res.* **105**, 21471–21481 (2000).
5. Hirshmann, M. M., Aubaud, C. & Withers, A. C. *Earth Planet. Sci. Lett.* **236**, 167–181 (2005).
6. Korenaga, J. in *Archean Geodynamics and Environments* (eds Benn, K., Mareschal, J.-C. & Condie, K.) 7–32 (Am. Geophys. Un., Washington DC, 2006).
7. Karato, S. *Nature* **347**, 272–273 (1990).
8. Mackwell, S. J. & Kohlstedt, D. L. *J. Geophys. Res.* **95**, 5079–5088 (1990).
9. Evans, R. L. *et al.* *Nature* **437**, 249–252 (2005).
10. Huang, X., Xu, Y. & Karato, S. *Nature* **434**, 746–749 (2005).
11. Bell, D. R., Rossman, G. R. & Moore, R. O. *J. Petrol.* **45**, 1539–1564 (2004).

Correction

In the News & Views article "Greenland's ice on the scales" by Tavi Murray (*Nature* **443**, 277–278; 2006), the figures given in mass units for Greenland's ice loss are 1,000 times too small. The correct statement is that the Greenland ice sheet lost between 192 billion and 258 billion — not million — tonnes each year between April 2002 and April 2006. The equivalent volume figures given in the article — 212–284 km³ — are correct. Our thanks to Mary Whitfield and her students at the Edmonds Community College, Lynnwood, Washington state, for pointing this error out.

BRIEF COMMUNICATIONS

Skull morphology of giant terror birds

These monstrous birds were probably more agile and less portly than previously thought.

The phorusrhacids ('terror birds') are an extinct lineage that includes the largest birds known^{1–3}. Reconstructions of these Cenozoic carnivores have consistently highlighted a very high beak, round orbits and vaulted braincase, although minimal information is actually available for the skull of the largest species^{1–5}. An important new fossil of a gigantic avian skull has been discovered from the middle Miocene of Patagonia (Comallo, Argentina), which reveals significant differences between the skulls of large and small phorusrhacids. We conclude that reconstructions of the skull of gigantic phorusrhacids on the basis of their smaller relatives are unwarranted, and that the long-established correlation between their corpulence and reduced cursorial agility needs to be re-evaluated.

The enormous skull of specimen BAR 3877-11 is virtually complete. Its length (tip of rostrum to sagittal nuchal crest) is about 716 mm, making it the largest known avian skull. It is triangular in dorsal view, with a very long and hooked rostrum (exceeding half the length of the skull) and a dorsoventrally compressed caudal portion. The small external nares and larger antorbital fenestra are, respectively, rectangular and quadrangular in shape. The orbits are low and subrectangular, and their ventral margin is formed by a stout and very tall jugal bar. The skull roof behind the orbits is flat and distinctly scarred by the extensive development of the temporal musculature. The occipital table is wide and low, giving the skull a rectangular appearance in occipital view.

The enormous size, laterally compressed and strongly hooked rostrum, and convex culmen support the assignation of BAR 3877-11 within the phorusrhacids^{3,6}. Its skull resembles that of the poorly preserved and slightly smaller *Devincenzia pozzii*³. However, the discovery of BAR 3877-11 reveals that these birds show significant differences with respect to their smaller and more gracile relatives⁷. Particularly evident are the proportionally much lower and longer rostrum, rectangular orbit, notable height of the jugal bar, flat cranial roof, and the proportionally low and rectangular-shaped occipital table characteristic of the skull of gigantic phorusrhacids such as BAR 3877-11 (Fig. 1).

Before the discovery of BAR 3877-11, knowledge of the cranial morphology of large-bodied phorusrhacids — those with skulls exceeding 600 mm in length — was limited to the fragmentary *D. pozzii*^{3,8} and a sketch of the

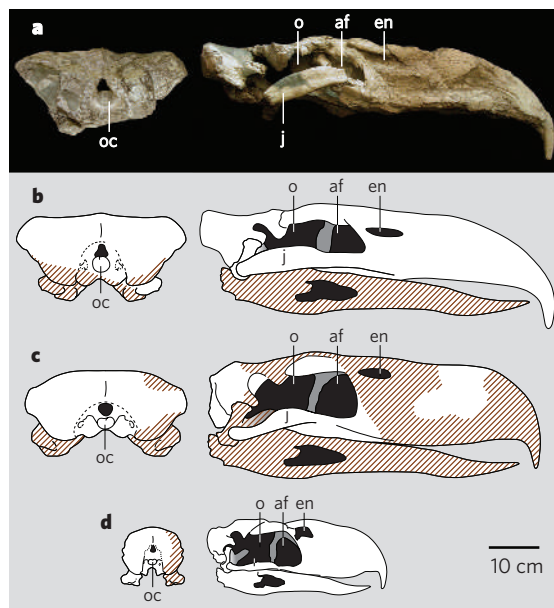


Figure 1 | Skulls of terror birds. a, Photograph of the skull (specimen BAR 3877-11; Museo Asociación Paleontológica Bariloche, Río Negro Province, Argentina). b–d, Reconstructions of BAR 3877-11 (b), *Devincenzia pozzii* (c) and *Patagornis marshi* (d) in occipital (left) and lateral (right) views. The skull of BAR 3877-11 was found by Guillermo Aguirre-Zabala, adjacent to a tarsometatarsus and fragments of pedal phalanges. Abbreviations: af, antorbital fenestra; en, external nares; j, jugal; o, orbit; oc, occipital condyle. (Figure courtesy of S. Abramowicz.)

skull of *Phorusrhacos longissimus*, which was destroyed during collection⁹. Perhaps influenced by the latter, a sketch largely inspired by the morphology of the much smaller *Patagornis marshi*^{3,10} (skull length about 350 mm), the skull of gigantic phorusrhacids has been consistently interpreted as a scaled version of those of their much smaller and better-known relatives^{1–3,5}. However, the discovery of BAR 3877-11 and our reinterpretation of the skull of *D. pozzii* suggest that this widely accepted notion is incorrect.

BAR 3877-11 also bears on a long-standing view about the evolution of phorusrhacids. The broad range of sizes and inferred agility of phorusrhacids — from maximum standing heights similar to the 0.9-m red-legged seriama to birds that would have stood close to 3 m high — has been traditionally subdivided into groups of increased corpulence and proportionally reduced agility^{1–3,11–13}. On the basis of length measurements of the skull and tarsometatarsus, BAR 3877-11 is estimated to be 10% larger than the largest phorusrhacids previously known (*P. longissimus* and *Bronthornis burmeisteri*)³. However, the long (437 mm) and rather slender (midshaft width/length ratio of about 0.11) tarsometatarsus of BAR 3877-11 indicates that this bird may have been substantially swifter than the graviportal *B. burmeisteri*. Therefore, the discovery of BAR 3877-11 indicates that the traditional inverse

correlation between body size and cursorial agility^{1–3,11–13} needs to be reconsidered.

Luis M. Chiappe, Sara Bertelli

The Dinosaur Institute, Natural History Museum of Los Angeles County, Los Angeles, California 90007, USA

e-mail: chiappe@nhm.org

1. Feduccia, A. *The Origin and Evolution of Birds* (Yale Univ. Press, New Haven, 1999).
2. Marshall, L. G. *Sci. Am.* **270**, 90–95 (1994).
3. Alvarenga, H. & Höfling, E. *Pap. Avulsos Zool.* **43**, 55–91 (2003).
4. Kraglievich, L. *An. Mus. Hist. Nat. Montevideo* **3**, 323–353 (1932).
5. Gould, G. C. & Quitmyer, I. *Bull. Fla. Mus. Nat. Hist.* **45**, 201–229 (2005).
6. Livezey, B. C. *Trans. Zool. Soc. Lond.* **353**, 2077–2151 (1998).
7. Sinclair, W. J. & Farr, M. S. in *Reports of the Princeton University Expeditions to Patagonia, 1896–1899* (ed. Scott, W.) 157–191 (Princeton Univ. Press, New Haven, 1932).
8. Patterson, B. & Kraglievich, J. L. *Publ. Mus. Mun. Cienc. Nat. Trad. Mar del Plata* **1**, 1–52 (1960).
9. Ameghino, F. *Bol. Inst. Geogr. Argent.* **15**, 501–602 (1895).
10. Andrews, C. W. *Trans. Zool. Soc. Lond.* **15**, 55–86 (1899).
11. Tonni, E. P. *Nat. Hist. Mus. Los Angeles County* **330**, 105–114 (1980).
12. Tambussi, C. & Noriega, J. in *Contributions of Southern South America to Vertebrate Paleontology* (ed. Arratia, G.) 245–264 (Müncher Geowiss, München, 1996).
13. Alvarenga, H. *An. Acad. Bras. Cienc.* **54**, 697–712 (1982).

Received 12 August; accepted 28 September 2006.

Competing financial interests: declared none.
doi:10.1038/443929a

BRIEF COMMUNICATIONS ARISING online
www.nature.com/bca see Nature contents.

EVOLUTIONARY BIOLOGY

Sympatric plant speciation in islands?

Arising from: V. Savolainen *et al.* *Nature* **441**, 210–213 (2006)

Comparative studies of populations, particularly with the help of molecular markers, are necessary to understand the mechanisms of speciation in isolated oceanic archipelagos. Savolainen *et al.*¹ present comparative data on two endemic species of the palm genus *Howea* in Lord Howe Island, from which they infer that speciation was sympatric — that is, divergence had occurred in the absence of geographic isolation. However, the landscape of oceanic islands changes dramatically over time, with many disappearing under the sea after 6 million years or more, and Lord Howe Island is in a very late stage of its ontogeny. An alternative explanation, therefore, is that these two species did not evolve *in situ* but instead that they arose allopatrically after becoming geographically isolated at a time when the island was much larger and more diverse ecologically.

Aside from well-known mechanisms of chromosomal polyploidy, examples of homoploid sympatric speciation in plants are rare. Savolainen *et al.*¹ present interesting comparative morphological, geographical, ecological and molecular (AFLP, or amplified DNA-fragment length polymorphism) data regarding relationships between the two species of *Howea* palms endemic to Lord Howe Island. Considering that the distributions of the two taxa now overlap, the authors regard this as a good example of sympatric plant speciation. The difficulty with this interpretation is that it does not take into account island ontogeny and its possible impact on the two species. Sympatric

speciation in islands, therefore, may still not have been convincingly demonstrated.

Lord Howe Island has a geomorphology that suggests strong erosion and surface loss, which correlates with its advanced age of 6.4–6.9 Myr. Dating from molecular sequences reveal the genus *Howea* to have diverged from progenitors 4.5–5.5 Myr ago, which is compatible with the island's age. By the same methods, the two species are judged to be 0.5–1.9 Myr old, at which time the island would still have been much larger than today. Loss of terrestrial area is more rapid during the later stages of island ontogeny, when the sea covers the last parts of the land.

Many other examples exist of sympatric and closely related plant species in oceanic islands. Such a pattern is evident among six species of *Robinsonia* (Asteraceae) from Masatierra Island in the Juan Fernandez archipelago in Chile². My colleagues and I have estimated that this island, which is about 4 Myr old³, has lost up to 95% of its original surface area⁴. The pattern we now see can best be interpreted as refugial, with these species being packed together (sympatrically) in the remaining small surface area. The original mode of speciation would have been ecogeographical (allopatric).

An alternative hypothesis for the origin of the two endemic species of *Howea* on Lord Howe Island, therefore, is by allopatric speciation in different ecological (perhaps calcarenite) zones, when the island was much younger and much larger. The low degree of AFLP divergence between the two species of

Howea does not necessarily argue for sympatric speciation. Island congeners typically show very little genetic divergence in DNA sequences and isozymes⁵. An absence of broad surveys prohibits generalizations at this point, but the limited AFLP data available indicate that very little divergence may have occurred between closely related species in islands that have speciated allopatrically⁶.

Before jumping to conclusions about sympatric speciation in oceanic islands, we need to exercise caution and factor in our best estimate of island geomorphic and habitat change, which together have had a major impact on genetic and species diversity⁷.

Tod F. Stuessy

Department of Systematic and Evolutionary Botany, University of Vienna, 1030 Vienna, Austria
e-mail: tod.stuessy@univie.ac.at

1. Savolainen, V. *et al.* *Nature* **441**, 210–213 (2006).
2. Sanders, R. W., Stuessy, T. F., Marticorena, C. & Silva, O. M. *Opera Bot.* **92**, 195–215 (1987).
3. Stuessy, T. F., Foland, K. A., Sutter, J. F., Sanders, R. W. & Silva, O. M. *Science* **225**, 49–51 (1984).
4. Stuessy, T. F., Crawford, D. J., Marticorena, C. & Rodriguez, R. in *Evolution and Speciation of Island Plants* (eds Stuessy, T. F. & Ono, M.) 121–138 (Cambridge Univ. Press, 1998).
5. Crawford, D. J. & Stuessy, T. F. in *Evolution and Diversification of Land Plants* (eds Iwatsuki, K. & Raven, P. H.) 249–267 (Springer, Tokyo, 1997).
6. Parsons, Y. M. & Shaw, K. L. *Mol. Ecol.* **10**, 1765–1772 (2001).
7. Stuessy, T. F., Greimler, J. & Dirnböck, T. *Biol. Skrift.* **55**, 89–101 (2005).

doi:10.1038/nature05216

EVOLUTIONARY BIOLOGY

Savolainen *et al.* reply

Replying to: T. F. Stuessy *Nature* **443**, doi10.1038/nature05216 (2006)

Stuessy¹ questions our conclusions of sympatric speciation in a case study of palms on Lord Howe Island² and proposes an alternative hypothesis, whereby the two *Howea* species evolved allopatrically when the island was larger and less eroded. Stuessy also argues that low genetic divergence does not necessarily indicate speciation in sympatry¹. We agree that it is important not to jump to conclusions, but we have good estimates of the size and geological history of Lord Howe Island at the time of the speciation event^{3,4}, and both are fully compatible with sympatric speciation. Stuessy also misinterprets the results from our

AFLP (amplified DNA-fragment length polymorphism) genome scan: we did not assert that low AFLP divergence *per se* is evidence for sympatric speciation, but rather that the distribution of these genetic divergence values across the genome is strongly supportive of speciation with gene flow involving disruptive or divergent selection².

Evidence from bathymetry around the island indicates that its past extent has not exceeded 30 km long by 23 km wide³; today the island is roughly 12 km long and, on average, 1.5 km wide. The original height of the island is estimated to be 1,000 m, comparable with 875 m

for Mount Gower today³. The presence of an evenly distributed, wave-cut platform surrounding Lord Howe Island indicates that erosion has been mainly coastal and equal from all sides. Consequently, Quaternary calcarenite deposits, which created divergent ecological selection pressures conducive to *Howea* species divergence², have formed evenly around the island⁴; these are so closely intercalated with volcanic rocks that allopatric speciation due to ecogeographic isolation, as Stuessy proposes¹, is unrealistic.

The geomorphological history of Lord Howe Island must also be seen in the light of the breeding system of *Howea*. Even at its greatest extent, the island would have been small compared with the gene dispersal capacity of these woody, wind-pollinated outcrossers. Indeed, our analysis of molecular variance revealed extremely homogeneous population structure in both palms, which is compatible

with widespread gene flow². It is therefore difficult to imagine how geographic barriers to gene exchange would have formed on the island. The interlinked ecological differences of both species are still clearly visible — that is, in phenology and soil pH (ref. 2), which are both properties that are interdependent in other plants⁵.

Our AFLP genome scan is consistent with speciation in the face of gene flow, involving divergent selection on a limited number of genes. The expected distribution of interspecific divergence (F_{ST}) for species that evolved in allopatry is different from that expected for species that diverged with gene flow^{6–8}. It is the shape of the distribution of F_{ST} in the genome that matters: F_{ST} is expected to follow an L-shaped

distribution for sympatric divergence, as in our case study², whereas genetic differences accumulate throughout the genome in allopatry and lead to a bell-shaped distribution or a distribution that is biased towards large values of F_{ST} . Regardless of Stuessy's concerns, the question now is not whether sympatric speciation exists, but how common it might be^{9,10}.

Vincent Savolainen*, **Christian Lexer***,
Marie-Charlotte Anstett†, **Ian Hutton‡**,
J. J. Clarkson*, **M. V. Norup*S**, **M. P. Powell***,
D. Springate*, **N. Salamin||**, **William J. Baker***

*Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, UK

e-mail: v.savolainen@rbgkew.org.uk

†Centre for Evolutionary and Functional Ecology, UMR 5175, 34293 Montpellier cedex 5, France

‡PO Box 157, Lord Howe Island, New South Wales 2898, Australia

SDepartment of Systematic Botany, University of Aarhus, 8000 Aarhus, Denmark

||Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland

1. Stuessy, T. *Nature* **443**, doi:10.1038/nature05216 (2006).
2. Savolainen, V. et al. *Nature* **441**, 210–213 (2006).
3. McDougall, I. et al. *J. Geol. Soc. Austr.* **28**, 155–176 (1981).
4. Brooke, B. P. et al. *Quat. Sci. Rev.* **22**, 859–880 (2003).
5. Ollerton, J. *Heredity* **95**, 181–182 (2005).
6. Luikart, G. et al. *Nature Rev. Genet.* **4**, 981–994 (2003).
7. Via, S. *Trends Ecol. Evol.* **16**, 381–390 (2001).
8. Wu, C. I. *J. Evol. Biol.* **14**, 851–865 (2001).
9. Pennisi, E. *Science* **311**, 1372–1374 (2006).
10. Ortiz-Barrientos, D. & Rieseberg, L. H. *Heredity* **97**, 2–3 (2006).

doi:10.1038/nature05217

Insights into social insects from the genome of the honeybee *Apis mellifera*

The Honeybee Genome Sequencing Consortium*

Here we report the genome sequence of the honeybee *Apis mellifera*, a key model for social behaviour and essential to global ecology through pollination. Compared with other sequenced insect genomes, the *A. mellifera* genome has high A+T and CpG contents, lacks major transposon families, evolves more slowly, and is more similar to vertebrates for circadian rhythm, RNA interference and DNA methylation genes, among others. Furthermore, *A. mellifera* has fewer genes for innate immunity, detoxification enzymes, cuticle-forming proteins and gustatory receptors, more genes for odorant receptors, and novel genes for nectar and pollen utilization, consistent with its ecology and social organization. Compared to *Drosophila*, genes in early developmental pathways differ in *Apis*, whereas similarities exist for functions that differ markedly, such as sex determination, brain function and behaviour. Population genetics suggests a novel African origin for the species *A. mellifera* and insights into whether Africanized bees spread throughout the New World via hybridization or displacement.

The western honeybee, *Apis mellifera*, is a striking creature, one of relatively few species for which evolution culminated in advanced society¹. In 'eusocial' insect colonies, populations are differentiated into queens that produce offspring and non-reproductive altruistic workers that gather and process food, care for young, build nests and defend colonies. Remarkably, these two castes, both highly derived relative to solitary insects, develop from the same genome.

Social evolution endowed honeybees with impressive traits^{2,3}. Differentiation into queens and workers is through nutritionally based, hormone-mediated, programmes of gene expression⁴ yielding dramatic distinctions in morphology, physiology and behaviour. Queens, typically one per colony, have ten times the lifespan of workers, typically 1 to 2 yr⁵, lay up to 2,000 eggs per day, and store sperm for years without losing viability. Workers, numbering tens of thousands per colony, display sophisticated cognitive abilities, despite a brain containing only one million neurons⁶. This is five orders of magnitude less than the human brain and only four times greater than *Drosophila*, which has a far simpler behavioural repertoire. Workers learn to associate a flower's colour, shape, scent, or location with a food reward⁷, increasing foraging efficiency. They communicate new food discoveries with 'dance language', originally deciphered by von Frisch⁸, the only non-primate symbolic language. Recent studies revealed that honeybees can learn abstract concepts such as 'same' and 'different'⁹.

The infamous African 'killer' bees, *Apis mellifera scutellata*, the queens of which were introduced to Brazil in 1956¹⁰, are known for intense stinging activity during nest defence, and pose human health problems. The African bees' spread throughout the New World is a spectacular example of biological invasion. Although it was one of the first biological invasions to be studied with molecular tools¹¹, our understanding of its genetic basis has been controversial.

This array of fascinating features, as well as amenability to molecular, genetic, neural, ecological and social manipulation¹², led to selection of the honeybee for genome sequencing by the National Human Genome Research Institute, National Institutes of Health (NHGRI, NIH)¹³. The United States Department of Agriculture (USDA) also supported the project because of the paramount importance of pollination to human nutrition and the environment¹⁴. And, of

course, humans and other animals have valued honey since prehistoric times.

Honeybees belong to the insect order Hymenoptera, which includes 100,000 species of sawflies, wasps, ants and bees. Hymenoptera exhibit haplodiploid sex determination, where males arise from unfertilized haploid eggs and females arise from fertilized diploid eggs. Haplodiploid-induced asymmetries in relatedness between offspring and sisters have long been thought to be involved in the evolution or maintenance of eusociality in the Hymenoptera^{15,16}, but other life history traits also promote social evolution¹⁷, and there are divergent perspectives on this issue at the present time^{1,18}. Haplodiploidy has distinct sex-determination mechanisms compared with other organisms because Hymenoptera lack sex chromosomes¹⁹.

Hymenoptera is one of 11 orders of holometabolous (undergo a metamorphic moult) insects. All completed insect genome sequences have thus far been confined to Holometabola^{20–26}; phylogenetic relationships of these and related arthropods are in Fig. 1. Honeybees diverged from Diptera and Lepidoptera 300 million years ago, whereas the last common ancestor with humans was 600 million years ago²⁷. The genus *Apis* is an ancient lineage of bees that evolved in tropical Eurasia²⁸ and migrated north and west, reaching Europe by the end of the Pleistocene epoch, 10,000 yr ago. The origin of *A. mellifera* has been suggested as Asia²⁸, the Middle East²⁹, or Africa^{2,30}. From there, humans carried them worldwide because of their ability to make honey²⁸.

The *A. mellifera* genome has novel characteristics and provides fascinating insights into honeybee biology. Some main findings are:

- The *A. mellifera* genome is distinguished from other sequenced insect genomes by high A+T content, greater spatial heterogeneity of A+T content, high CpG content, and an absence of most major families of transposons.

- The honeybee genome evolved more slowly than that of the fruitfly and malaria mosquito.

- The *A. mellifera* genome shows greater similarities to vertebrate genomes than *Drosophila* and *Anopheles* genomes for genes involved in circadian rhythms, RNA interference (RNAi) and DNA methylation, among others.

*Lists of participants and affiliations appear at the end of the paper.

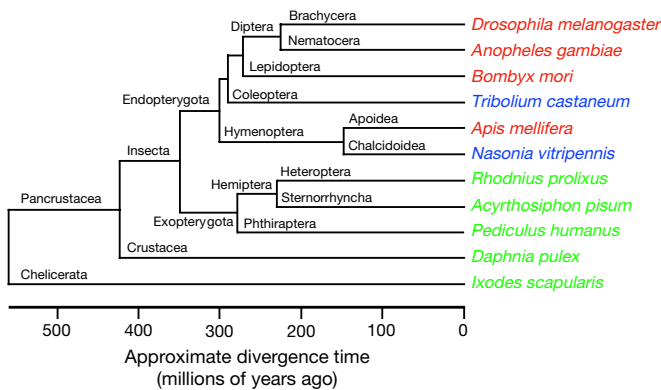


Figure 1 | Evolutionary relationships. Evolutionary relationships of *Apis mellifera*, other insects and related arthropods for which the genome sequence has been published (red), is in draft assembly form (blue), or is approved for sequencing (green), with approximate divergence times^{161,180}. Recent work suggests that the Hymenoptera are basal to the Coleoptera in the Endopterygota (also known as Holometabola)^{76,152}.

- *Apis mellifera* has fewer genes than *Drosophila* and *Anopheles* for innate immunity, detoxification enzymes, cuticle-forming proteins and gustatory receptors, but more genes for odorant receptors, and novel genes for nectar and pollen utilization. This is consistent with honeybee ecology and social organization.
- Genes encoding the major royal jelly protein family—nine genes evolved from one ancient *yellow* gene—involved in queen and brood nursing, exemplify genes gaining new functions during the evolution of sociality.
- Novel microRNAs (miRNAs) were detected and shown to have caste- and stage-specific expression, suggesting a role in social diversification.
- Key elements in early developmental pathways differ between *Apis* and *Drosophila*, indicating that these evolved after the lineages separated.
- The honeybee shows similarities to *Drosophila* for functions that differ markedly, such as sex determination, brain function and behaviour.
- Population genetic analyses using new genome-based single-nucleotide polymorphisms (SNPs) support a hypothesis involving an African origin for the species *A. mellifera* and provide new insights into the spread of Africanized 'killer' bees. *A. m. scutellata* alleles have largely replaced those from one previously dominant subspecies, *A. m. ligustica*, whereas *A. m. mellifera* genotypes were essentially unchanged³⁰.

Genome sequencing and assembly

The honeybee genome was sequenced using DNA from multiple drones derived from a single, slightly inbred queen (DH4 strain; Bee Weaver Apiaries, Inc.). 2.7 million whole-genome shotgun reads

(Supplementary Tables 2 and 3) were assembled using the Atlas software³¹ and built into chromosomes using a microsatellite marker linkage map^{32,33}. Analysis of initial assemblies showed that regions with high A+T composition were under-represented in libraries. Previously as much as 30% of the genome was in an (A+T)-rich shoulder in density gradients³⁴. Thus, DNA was fractionated by CsCl-bisbenzimidazole density gradient centrifugation³⁵ and additional shotgun libraries were generated from DNA with >70% A+T composition. Batches of 200,000 reads of (A+T)-rich DNA were generated and new genome assemblies were reassessed. After four such batches were added to the assembly, representing about 30% of the data, coverage of the (A+T)-rich regions had improved to 6-fold and the N50 (N50 is the contig size where 50% of the genome sequence is in contigs of size N50 or larger) of contigs had doubled to over 30 kilobases (kb) (Supplementary Fig 2 and 3), statistics adequate for gene predictions and analysis. In total, 1.8 gigabases (Gb) were assembled, $\times 7.5$ coverage of the (clonable) 236 megabase (Mb) honeybee genome (Table 1, Supplementary Tables 2 and 3). Further details are available in Supplementary Information.

Several assemblies were produced (Supplementary Table 1) and are available on the Baylor College of Medicine Human Genome Sequencing Center (BCM-HGSC) website³⁶, with statistics displayed in Table 1. Versions 1.x were initial assemblies deficient in (A+T)-rich regions. Version2 (January 2005) was the first to contain full (A+T)-rich read enrichment. The genome size and coverage of this assembly did not change for later assemblies. Version2 lacks highly repeated sequences, which has little impact on gene predictions, and used the 2005 version of the microsatellite linkage map (from M. Solignac) to build chromosomes. In version3, highly repeated sequences were added, which increased the N50 (see below) for contigs by 15% and for scaffolds by 6%. With the current (March 2006) version4 assembly, the most recent genetic map (AmelMap³³) was used to place sequence on chromosomes, increasing the amount of mapped sequence by 10%.

Assemblies were tested against honeybee data sets (expressed sequence tags (ESTs), complementary DNAs, microsatellite markers and sequenced bacterial artificial chromosome (BAC) clones) for quality and completeness (Supplementary Table 3). The quality of version3 and version4 was tested by alignment of full-length cDNAs (55 total). All alignments showed correct order and orientation of exons. Although a limited set of cDNAs, it nevertheless seems that misassembly of genic regions is rare. In version4, 99% of 2,032 markers and 98% of 3,136 ESTs were represented, providing evidence for completeness. When aligned to BACs that had been independently sequenced, 23 out of 27 BACs showed >94% coverage, whereas four were >80% but were complex owing to repeated sequences. The estimated clonable genome size is thus taken as 231 Mb/98% = 236 Mb. Flow cytometry estimates with nuclei from brains of 45 workers and 12 drones gave a haploid genome size of 262 ± 1 Mb. The difference between drones (264 ± 2) and workers (261 ± 1) was not significant ($P \gg 0.05$). Compared to the assembly, we estimate unassembled or unclonable (for example, paracentromeric (A+T)-rich, see below) sequences at 26 Mb.

Table 1 | Genome assembly statistics

Version	Reads (million)	Bases (Mb)	N50 contigs (kb)*	N50 scaffolds (kb)*	Total contigs (Mb)	Span contigs (Mb)	Coverage by reads (%)	Genome size (Mb)†	Coverage‡	Mapped (oriented) (Mb)	Mapped (not oriented) (Mb)	Unmapped (Mb)
1.0	2.0	1,251	22	236	176	182	>96	~212	~6×	81 (38%)	32 (15%)	99 (47%)
1.1	2.0	1,251	19	223	198	206	>96	210	6.0×	81 (39%)	29 (14%)	100 (47%)
1.2	2.0	1,374	23	260	208	213	>96	219	6.3×	101 (47%)	29 (14%)	79 (39%)
2.0	2.7	1,776	36	343	225	229	>96	237	7.5×	132 (58%)	28 (12%)	77 (30%)
3.0	2.7	1,776	41	362	231	235	>96	238	7.5×	138 (59%)	28 (12%)	72 (29%)
4.0	2.7	1,776	41	362	231	235	>96	236	7.5×	152 (65%)	34 (14%)	49 (21%)

* The N50 size is the length such that 50% of the assembled genome lies in blocks of the N50 size or longer.

† Genome size = $\sum(\text{contig lengths})/(\text{completeness})$.

‡ Coverage = bases/genome size.

Data are available at <http://www.hgsc.bcm.tmc.edu/projects/honeybee>.

Sets of gene predictions were produced using version2 (see below and Supplementary Information). Analyses proceeded using versions 2, 3 and 4, depending on the particular characteristic under study and timing of the assembly.

Genetic and physical maps and chromosome organization. Development of linkage maps was facilitated by a high recombination rate, 19 cM Mb^{-1} , several-fold greater than in any other multicellular eukaryotes^{32,37,38}. The linkage map AmelMap3 has more than 2,000 microsatellite markers. The average distance between markers is 2.1 cM; all intervals are shorter than 10 cM. The high density of this map suggests that little information is missing in the assembly. Scaffolds were organized along chromosomes according to this map. Sixty-four per cent (151 Mb) of scaffolds contain at least two markers with non-null distances and could be oriented on chromosomes, whereas another 15% (35 Mb) contained one marker and could be placed but not oriented. Thus, 79% (186 Mb) of the genome is placed on chromosomes.

The honeybee karyotype (Fig. 2) is based on measurements of morphologically distinct features of chromosomes in 74 well-spread, 4,6-diamidino-2-phenylindole (DAPI)-stained haploid chromosomes prepared from testis of drone pupae of the sequenced strain. Measurements of centromere positions, overall chromosome length, position of ribosomal organizer regions, and position and extent of (A+T) sequence-rich (DAPI-positive bands) are in Supplementary Table 4. The total length of the 16-chromosome haploid complement at meiotic pre-metaphase is 30 μm . Of this, 36% (11 μm) is (A+T)-rich, DAPI-positive sequence surrounding each centromere. The chromosomes are numbered according to the genetic map, which orders them roughly by length, ranging in size from 3.5 μm for metacentric chromosome 1 to 1.2 μm for chromosome 16. All chromosomes (plus nuclear organizing regions containing rDNA repeats on chromosomes 6 and 12) have been identified by BAC fluorescence *in situ* hybridization (FISH) using one or more BACs containing microsatellites used in the genetic mapping effort. This karyotype largely agrees with previous efforts³⁹; however, it is not possible to pair unequivocally the several similar-sized chromosomes in these two karyotypes.

Manual superscaffolding of chromosomes 13, 14, 15 and 16. Additional evaluation of completeness of version4 was performed using 'manual superscaffolding' to bridge gaps between mapped scaffolds on the four smallest chromosomes. The requirement for two unconflicted mate pairs to link contigs was relaxed and all additional sources of evidence were used. These included overlaps of contigs

that were not merged because of haplotype divergence, single-mate-pair links between contigs, extensions of contigs using trimmed parts of reads yielding novel overlaps, resolution of conflicting mate pairs, and cDNAs and confident gene models bridging inter-scaffold gaps. Attempts were made to cross the 22 remaining inter-superscaffold gaps by polymerase chain reaction (PCR), with five being successful. This effort reduced chromosomes 13–16 (17% of the mapped genome) from 21, 25, 42 and 21 mapped scaffolds to 4, 5, 6 and 5 superscaffolds, respectively, and incorporated 121 unmapped scaffolds totalling 1.8 Mb for an increase in length of 5.5% for these chromosomes (superscaffolds are available from BeeBase).

The resultant superscaffolds extend from the mapped location of the centromere—although centromeric sequences were not discerned—to the TTAGG telomeric repeats of the distal telomeres (see below). Comparison with the genetic map suggests that the remaining inter-superscaffold gaps are not extensive. Only two misassemblies of 146-kb and 65-kb sections of scaffolds, as well as several minor misassemblies of 2–8-kb contigs, were discovered in this 17% of the mapped assembly. This manual effort provides additional support for the near completeness of the assembly for the euchromatic regions of the genome.

Genome organization

Although the *A. mellifera* genome is not the first sequenced insect genome, a number of characteristics of genome organization distinguish it. Described below, the honeybee provides new diversity in genome structure.

Sequence characteristics. Animal genomes are a mosaic of G+C-content domains, with homogeneous G+C composition within domains, but widely variable G+C composition between domains. In all animals studied, including the honeybee, the distribution of G+C-content domain lengths follows a power law distribution (for example, ref. 40). The G+C-content domains in the honeybee genome, as in other genomes, do not have a characteristic length; rather, there is an abundance of short segments and only a small number of long ones. Comparison of G+C-content domain lengths in various genomes (Supplementary Fig. 4) shows that honeybee domains are shorter than in two dipterans, which are shorter than chicken and human. The honeybee genome is more (A+T)-rich than other sequenced insect genomes (67% A+T in honeybee, compared with 58% in *Drosophila melanogaster* and 56% in *Anopheles gambiae*, Fig. 3a).

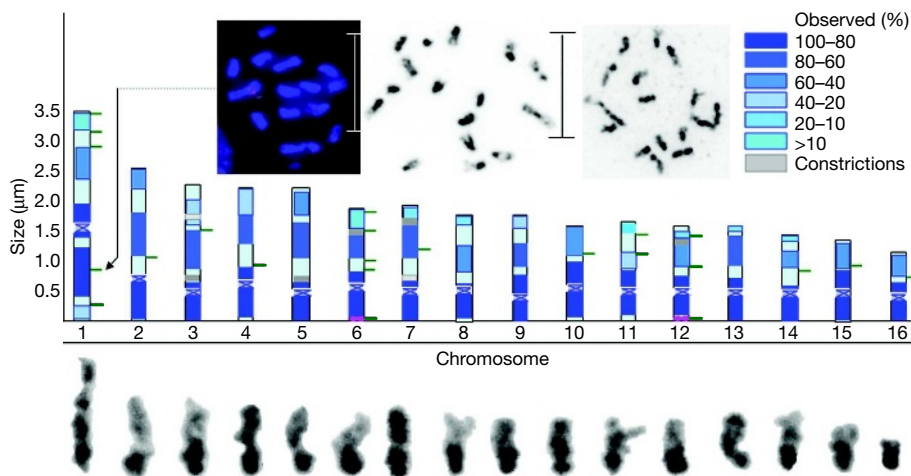


Figure 2 | Chromosomal spreads, ideogram and karyotype of *Apis mellifera*. The ideogram (in blue) shows average chromosome lengths, positions and sizes of DAPI-positive (heterochromatin) bands. The percentage of heterochromatin reflects the time of appearance of heterochromatic bands (100% observed in all preparations; lower

percentages seen only in early prophase spreads). Lines to the right of chromosomes represent BACs shown by FISH to bind in relative order and positions predicted from the genetic and physical maps. Binding sites of rDNA probes (distal short arms of chromosomes 6 and 12) are shown in red. The karyotype (below the ideogram) is based on the rightmost spread.

Consistent with an (A+T)-rich genome, honeybee genes occur more frequently in (A+T)-rich domains compared with other species (Fig. 3b and Supplementary Fig. 5). The mean G+C content of domains in which honeybee genes are located is 29%, compared with 47% for *A. gambiae* and 44% for *D. melanogaster*. Furthermore, genes are not distributed evenly throughout the genome, but show a tendency to appear in (A+T)-rich regions of the honeybee genome (C. Elsik *et al.*, personal communication).

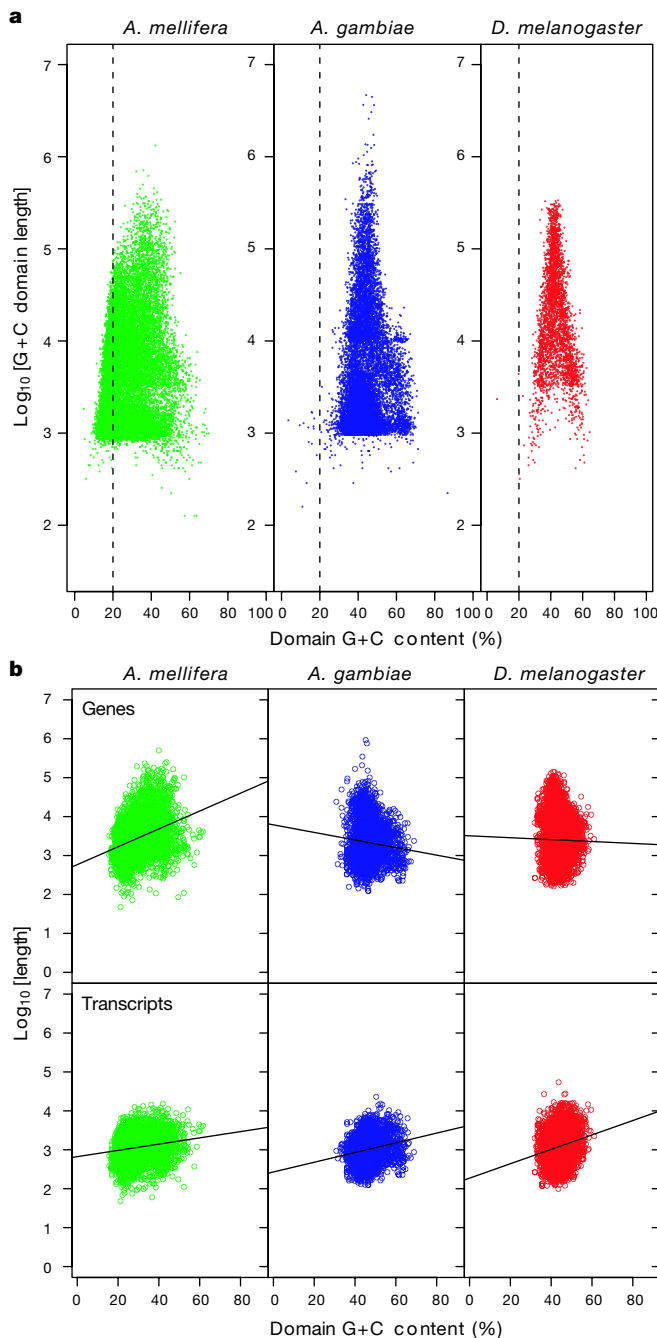


Figure 3 | Base composition in *Apis*, *Drosophila* and *Anopheles*. **a**, G+C-content domain length versus G+C percentage in *A. mellifera* (green), *A. gambiae* (blue) and *D. melanogaster* (red). The dashed line at 20% G+C content indicates the large number of low-G+C domains in *A. mellifera*. **b**, Gene length (top) and transcript length (bottom) versus G+C percentage of G+C content domains in which genes are embedded. Gene length was computed as the genomic distance from start to stop codon of the longest splice variant of each gene. Transcript length was computed as the distance between start and stop codon on the transcript sequence.

Gene length in the honeybee shows a striking relation to G+C content compared with other insects (Fig. 3b). Gene length (exons plus introns) increases with G+C content in honeybee ($R^2 = 0.135$, $P < 2.2 \times 10^{-16}$), but decreases slightly with G+C content ($R^2 = 0.009$, $P = 1.1 \times 10^{-10}$) in *A. gambiae*, and is not significantly related to G+C content ($R^2 = 0.0002$, $P = 0.079$) in *D. melanogaster*. However, transcript length (exons only) has a significant ($P < 2.2 \times 10^{-16}$) but weak positive correlation with G+C content in all three insects (*A. mellifera* $R^2 = 0.067$, *A. gambiae* $R^2 = 0.037$, *D. melanogaster* $R^2 = 0.033$). These relationships in honeybee indicate that total intron length increases with G+C content, in contrast with vertebrates in which intron length decreases with G+C content⁴¹. The relationship between gene length and G+C content is unlikely to be the result of annotation bias. First, this analysis included only genes that were flanked on both sides by other genes on the same scaffold, avoiding genes with exons on different scaffolds. Second, honeybee gene models were a consensus of five gene sets (see below), and were associated with probabilistic confidence scores. The relationship between gene length and G+C content did not change when gene models with less than 90% probability were removed from the analysis. Thus the honeybee genome is unusual among insects and opposite of vertebrate genomes in having long genes in (G+C)-rich regions.

Among dinucleotides, CpG is over-represented (1.67-fold) compared with the expectation from mononucleotide frequencies. Such an excess contrasts with other eukaryotic genomes where CpG is under-represented or at most close to the expected value (the highest CpG content so far is 1.15-fold in the genome of *Cyanidioschyzon merolae*⁴²). Genomes where CpGs are the target of cytosine methylases (such as plants or vertebrates) show a CpG deficit (Supplementary Table 5). The high CpG content in the honeybee is intriguing because honeybee is the first protostome shown to possess a full complement of functionally active vertebrate-like DNA methyltransferases, including an active CpG methyltransferase (see below). Methylated cytosines are known to be hypermutable⁴³, and MeC→T mutations are expected, driving base composition towards A+T richness. However, the impact of cytosine methylation on mono- and dinucleotide composition in the honeybee is not clear.

Telomeres. The 15 acrocentric honeybee chromosomes have a distal telomere on their long arm and a proximal telomere on their short arm, whereas the large metacentric chromosome 1 is presumably a centromeric fusion of two acrocentric chromosomes with two distal telomeres and loss of the proximal telomeres³⁹ (Fig. 2 and Supplementary Table 4). We built all 17 distal telomeres⁴⁴. Twelve were already present at the ends of terminally mapped scaffolds, whereas the remainder were assembled by manually superscaffolding outwards from the terminally mapped scaffolds. After 1–7 kb of unique sequence beyond the last gene on each *Apis* chromosome, each distal telomere has a 3–4-kb subtelomeric region, showing 70–92% sequence identity between all 17 telomeres. This is followed by the expected TTAGG⁴⁵ or variant telomeric repeats of at least several kilobases. The canonical organization of these distal telomeres (Fig. 4) makes them the simplest and most consistently constituted telomeres known in insects. In contrast, *Bombyx mori* telomeres are complicated by the insertion of numerous non-long terminal repeat (LTR) retrotransposons⁴⁶, and dipteran telomeres are unconventional in having no TTAGG repeats: *Drosophila* telomeres consist entirely of multiple non-LTR retrotransposon inserts of the het-A and TART families^{47–49}, and the telomeres of *A. gambiae*⁵⁰ and *Chironomus* midges⁵¹ consist of many complex tandem repeats, maintained by recombination.

We have been unable to build the 15 proximal telomeres, which also appear by FISH to have the TTAGG repeats⁴⁵, much as we have been unable to assemble the 16 centromeres, although their approximate map locations are known⁵². They are composed of highly repetitive tandem sequences of two major kinds: the 176-bp *AluI* repeat at proximal telomeres and the 547-bp *AvaI* repeat in centromeres³⁹.

These two repeats constitute 2% and 1% of the genome respectively, yet in the current assembly only a few short unmapped scaffolds contain them, whereas the vast majority remain in the unassembled-repeat-read data set. We conclude that the distal and proximal telomeres differ in their subtelomeric region sequences, even if both have terminal TTAGG telomere repeats. This subtelomeric difference might be important for the 15 acrocentric chromosomes, where in the Rabl configuration after mitosis the distal telomeres are sequestered on the nuclear envelope at one pole of the nucleus opposite the telophase centromeric cluster, as well as in the chromosome bouquet formation during meiotic prophase⁵³. To allow this Rabl configuration the proximal telomeres must be distinct, most likely in their subtelomeric regions, and hence remain with the centromeres. The apparent absence of *AluI* repeats from the single meta-centric centromere region is consistent with this interpretation³⁹.

Transposable and retrotransposable elements. Almost all transposons identified are members of the *mariner* family^{54–56}, widespread in arthropods and other animals, with relatives in plants, protists and bacteria⁵⁷ (Table 2 and Supplementary Information). They range in age from the relatively young *AmMar1* (refs 58, 59) with many nearly intact copies, to the ancient *AmMar5* and *AmMar6* with only a few highly degraded copies remaining. Other common types of transposons are largely absent from the honeybee genome (Supplementary Information).

There is little evidence for active retrotransposable elements, but the genome once harboured many retrotransposons. There are remnants of several LTR retroviral-like retrotransposons; for example, 15 partial and highly degraded copies of a copia-family sequence, 6 partial sequences encoding matches to the BEL12 element of *A. gambiae*⁶⁰, and 3 highly degraded copies of a DIRS retrotransposon⁶¹. There are also remnants of a further possible 11 LTR- and 7 non-LTR retrotransposons from a number of clades identified in *Drosophila*⁶². Separate manual assembly of nuclear rDNA units revealed that honeybees do contain active non-LTR retrotransposons of the R2 family, although the R1 lineage seems to be absent⁶³. There are, however, at least five short degraded copies of an R1-like element most similar to the R1Bm element of *B. mori*⁶⁴, so R1 elements were also once present. This relative lack of retrotransposons is a surprising finding in light of high diversity, copy number and activity of both LTR and non-LTR retrotransposons in other insect genomes. Possibly retrotransposons are too disruptive to a genome that is completely exposed to selection in haploid drones every generation,

Table 2 | *Mariner* family transposons

Name	Copy number*	Consensus length (bp)	Divergence from consensus (%)
<i>AmMar1</i>	360	1,287	4–5
<i>AmMar2</i>	100	1,284	6–9
<i>AmMar3</i>	390	1,304	8–10
<i>AmMar4</i>	80	1,310	3–15†
<i>AmMar5</i>	70	900‡	1–18†
<i>AmMar6</i>	130	600‡	14–20

* Copy number is approximate, and for assembled genome only.

† Several very similar copies are embedded within a longer recently duplicated sequence.

‡ These consensus sequences are not full-length.

with the rDNA being a relative ‘safe haven’, although retrotransposons are known from other haplodiploid Hymenoptera^{65–68}.

It is apparent that this is an unusual genome in having few transposons and retrotransposons, at least in assembled sequences, almost all of which are members of the *mariner* family constituting 1% of the assembled genome. The vast majority of these *mariner* copies and the degraded retrotransposons are in short unmapped scaffolds, suggesting that they might reside in poorly assembled pericentromeric regions that might constitute heterochromatin, and that there might be additional copies within the unassembled centromeric regions, just as there are R2 elements in unassembled rDNA repeats. Genomic screens for highly repetitive sequences identified only three major classes that represent about 3% of the genome, mapping to the centromere and telomere by chromosome *in situ* analyses^{39,45}. This 3% portion of the genome does not assemble well and could harbour undetected transposons.

Gene content and the proteome

The gene list. Five gene lists, each produced using the version2 assembly, were combined to produce a master gene list: the Official Gene Set (OGS). The component gene sets were from NCBI, Ensembl, Softberry (Fgenesh), an evolutionarily conserved core set and a set based on *Drosophila* orthologues (Supplementary Information). The gene sets were merged using GLEAN, which uses latent class analysis to estimate accuracy and error rates for each source of gene evidence, and then uses these estimates to construct a consensus prediction based on patterns of agreement or disagreement observed between each evidence source⁶⁹. The five input gene sets and merged GLEAN set were compared to each other using FASTA⁷⁰ to assess accuracy and completeness in representing 395 protein sequences that had been manually curated but not used to

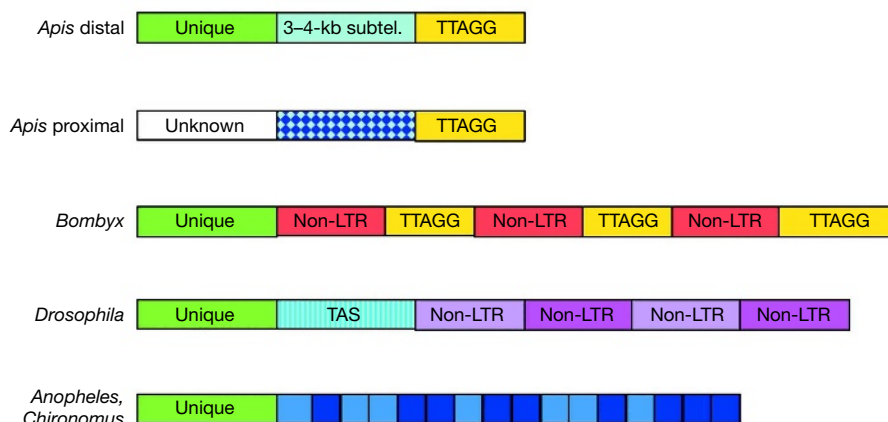


Figure 4 | Comparison of *Apis* telomeres with other insects. Organization of centromeric-proximal telomeres on the short arm of the 15 acrocentric chromosomes in *Apis* is hypothetical based on FISH studies. (The blue checked area represents the 176-bp tandem *AluI* repeats (see text).) *Bombyx*⁴⁶ and *Drosophila*⁴⁷ telomeres are based on one or two telomeres in

each species. Regions of non-LTR retrotransposons are indicated. For the *Drosophila* subtelomeric region (telomere-associated sequence repeats, TAS), shading indicates the presence of short (50–130 bp) repeat sequence blocks. Telomeres of *A. gambiae*⁵⁰ and *Chironomus* midges⁵¹ consist of complex tandem repeats, as indicated.

make gene predictions (a 'gold standard' collection of genes for evaluating different gene lists). The GLEAN set of 10,157 genes was superior by several measures (Supplementary Table 6) and was deemed to be the OGS. This set is considered as the list of genes that are based on experimental evidence. A second list of genes was constructed from gene predictions that were not strictly based on experimental evidence, the Official *ab initio* Gene Set (OAIGS), and comprised 15,500 Fgenesh gene models that did not overlap genes in the OGS. The annotation consortium manually annotated over 3,000 gene models using standard operating procedures developed by community members and BCM-HGSC. Most annotated gene models were from the OGS because most OAIGS models were not valid (see below). Global comparisons to other organisms relied on the OGS, whereas BLAST searches to identify honeybee orthologues of known genes (and families) used both the OGS and OAIGS.

Validation with whole-genome tiling array data. A complete transcription profile of the honeybee genome was analysed at high resolution using tiling arrays (M. P. Samanta *et al.*, personal communication). Seven million 36-mer oligonucleotide probes were selected to represent the *A. mellifera* genome, including all intergenic regions, then queried with pooled honeybee messenger RNA from multiple tissues and life stages (Supplementary Information). An established statistical technique was applied to determine whether an annotated gene was transcribed⁷¹ (Supplementary Information). Sixty-seven per cent of genes in the OGS were expressed; in contrast, only 5% of genes in the OAIGS were expressed. Similarly, less than 1% (4 out of 456) of OAIGS genes from 15 long (G+C)-rich chromosomal segments lacking OGS genes showed expression in another experiment. These results provide independent empirical support for the GLEAN-derived OGS. Transcriptional signals were also observed in 2,774 intergenic regions. Some of these novel transcripts are non-coding RNAs.

Validation via annotation of chromosome 15 and 16 superscaffolds. As another assessment of the quality of the OGS, manual annotation was performed. Because this is a time-consuming process, the two smallest chromosomes, chromosomes 15 and 16, were selected and their entire superscaffolds were manually annotated by carefully inspecting gene models from all gene prediction sets after BLAST comparison to known protein and EST/mRNA sequences (Supplementary Information). For chromosomes 15 and 16 respectively there were 720 and 337 gene models created, with 5 and 7 tRNAs, 5 and 14 pseudogenes with multiple frameshifts, and 71 and 62 splice variants; 188 and 116 gene models of the OGS were significantly corrected by adding/removing exons, adjusting splice sites and merging/splitting transcripts. Only 48 and 23 new protein-coding gene models were added to the OGS (less than 1%), including 40 and 15 that were previously supported by only Fgenesh *ab initio* gene models, consistent with the tiling array analysis. Only 56 and 21 transcripts on these two chromosomes (7%) were problematic due to assembly gaps or indels (insertions and/or deletions) affecting open reading frames, consistent with the view that the draft assembly is high quality in predicted genic regions. A putative function was assigned to 639 and 254 protein-coding genes for chromosomes 15 and 16, respectively.

MicroRNAs. Two computational surveys of the genome for miRNA-like sequences identified 65 candidate miRNAs (Supplementary Table 7), including orthologues of confirmed miRNAs in other organisms, and novel miRNA candidates comprised of micro-conserved elements. Seventeen candidates, including eight novel miRNAs, were selected for validation by RT-PCR.

Some putative miRNAs, including novel candidates, exhibited caste-, stage- and/or tissue-specific expression profiles, ranking among the top 10% of all tiling array signals (D. Weaver *et al.*, personal communication). For instance, two novel miRNAs (C5599F, C689F) were more strongly expressed in queen abdomen than in worker, are among the strongest tiling array signals, and are in human and *Drosophila* genomes. C689F is also one of the eight most

abundant miRNA-like transcripts in pupae. By contrast, novel miRNA C5560 displays differential developmental stage specificity and is more strongly expressed in worker pupae than queen, but is more abundant in all tissues and castes than any other putative miRNA tested. Thus, miRNAs may have a function in developmental regulation of social organization, and some new miRNA candidates may have diverse roles in regulating gene expression in other organisms.

Gene order in insects. As in vertebrates⁷², gene order in insects is under limited selection. Less than 7% of single-copy orthologues retain gene order in three-way genome comparisons of *Apis*, *Drosophila* and *Anopheles*. This fraction is about 10% in an *Apis*–*Drosophila* comparison, considerably lower than chicken–human (over 85%)⁷² although both genome pairs diverged approximately 300 million years ago. This discrepancy can be attributed to higher rates of genome evolution in insects (see ref. 73).

Beyond local gene arrangement, chromosome-level synteny can be established for species as divergent as *Drosophila* and *Anopheles*^{74,75} as both have five major chromosomal elements. However, only a few correspondences can be established between the 16 *Apis* chromosomes and *Drosophila* and *Anopheles* chromosomes. There are significantly more orthologues shared between *Drosophila* chromosomal arm 3R and *Apis* groups 5 and 15 as well as *Drosophila* chromosomal arm 2L and *Apis* group 4, pointing to a common origin⁷⁶, whereas no other chromosome regions between *Apis* and *Drosophila* show statistically significant enrichments of shared orthologues.

Orthology and rate of honeybee evolution. Single-copy orthologues conserved among many species are well suited for measuring differences in genome evolutionary rates. All genes were classified based on their homology to genes in other completely sequenced organisms, using only three vertebrates (human (*Homo sapiens*), chicken (*Gallus gallus*) and fish (*Tetraodon nigroviridis*)) in order to obtain a relatively balanced data set in terms of species divergence (Fig. 5 and Supplementary Tables 8–10). To characterize the conserved core of honeybee genes and estimate evolutionary rates from it, we identified single-copy orthologues likely to be present in all metazoans. In the OGS, approximately 30% fall into this category (for exact definitions, see Fig. 5). For the remainder (70%) of the predicted honeybee genes, orthologous relationships are more complex, involving many-to-many orthologues and nonuniformly occurring (patchy) orthologues, or not detectable. These genes are indicative of functional differences at various levels. For example, there are 1,052 honeybee genes with orthologues that are unique to the three insects considered (that is, they probably encode insect-specific phenotypes). This set compares to 3,816 human genes with orthologues only in the three vertebrates (see striped boxes in Fig. 5), suggesting a more complex gene pool coding for vertebrate-specific features.

Comparison of the 2,404 single-copy orthologues present in exactly one copy in each of the insects and in human revealed that the mean sequence identity between honeybee and human is considerably higher than that of fly and human (47.5% versus 44.5%, with *t*-test significance of 10^{-11} , see Fig. 6 and Supplementary Fig. 6) and also higher than between mosquito and human (46.6%). This indicates a considerably faster evolutionary rate in the fly lineage and points to the honeybee as the slowest evolving insect sequenced so far. Two other independent measures confirmed this observation, as various genome-based rate measures tend to give similar results⁷³. Patchy orthologous groups, defined as having at least one member lost in both insects and vertebrates (Fig. 5), contain non-essential genes that are relatively easy to lose during evolution. The high fraction of retained genes within this category in the honeybee shows that it lost fewer genes than either one of the Diptera species (Fig. 6; the higher number of losses in mosquito compared to fly might be due to lower quality of mosquito annotation). When comparing the fraction of conserved intron positions (Fig. 6) between insect genes and vertebrate orthologues, a massive loss of introns in Diptera becomes

apparent whereas the honeybee has kept almost 80% of the identifiable ancient introns since divergence from vertebrate ancestors. The higher number of introns in the honeybee (see Supplementary Table 11 and Supplementary Fig. 7 for more data on six genomes) also supports the conclusion that the honeybee is slower diverging and less derived than the other insects—fly and mosquito. This observation is reflected in other orthologue categories in Fig. 5. For example, the honeybee has fewer many-to-many orthologues than fly and mosquito, indicating slower evolution (but this might be partially due to a more stringent honeybee gene prediction). Taken together, analysis of honeybee orthologues is in accordance with a recent finding that early metazoans had many, intron-rich genes⁷⁷ that have been subsequently lost and/or simplified (in terms of intron numbers) in fast evolving lineages such as insects, wherein the honeybee seems to be less reductive than the two Diptera species—fly and mosquito.

There are several clear indications of honeybee-specific duplications implying unique functions. A striking example is farnesyl

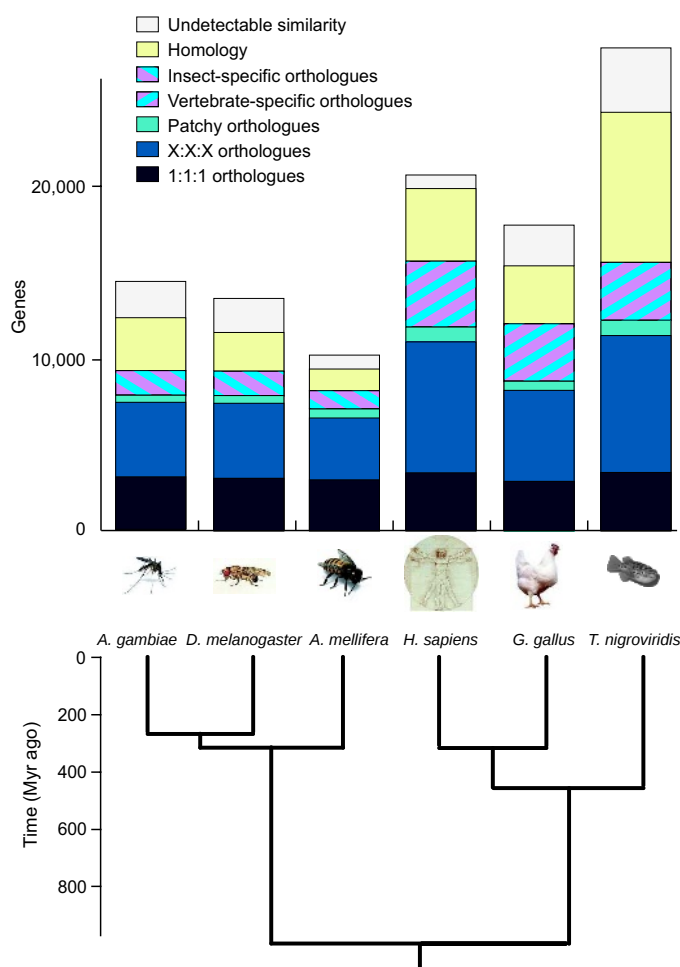


Figure 5 | Orthology assignment in insects and vertebrates. At the extremes, genes might be part of the metazoan core proteome (darkest band, bottom) or unique to a species with currently no counterpart in other organisms (white band, top). The striped boxes indicate insect- and vertebrate-specific genes and show that there are far fewer in insects. '1:1:1' indicates universal single-copy genes, but absence or duplication in a single genome is tolerated as we cannot exclude incomplete genomes or very recent duplications. This explains uneven numbers between species of these very conserved metazoan core genes. 'X:X:X' indicates any other orthologous group (miss in one species allowed), with X meaning one or more orthologues per species. 'Patchy' indicates other orthologues that are present in at least one insect and one vertebrate genome. 'Homology' indicates partial homology detected with $E < 10^{-6}$ but no orthology classified.

pyrophosphate synthase (FBgn0025373), known to be involved in lipid metabolism in fly, which occurs as a single copy in all other sequenced insects whereas there are seven copies in the honeybee. When strictly requiring single-copy orthologues in all five other metazoan genomes analysed, we find 60 genes that have been duplicated only in the honeybee (Supplementary Table 12). The fixation of these duplicates of otherwise single-copy orthologues is rare and has been shown to be associated with the emergence of species-specific functions⁷⁸. Some of these may be relevant to the solitary versus social lifestyle differences between these insect groups.

In addition to the presence and absence of genes and gene families, the expansion and reduction of the latter indicates, at a global level, change of functionality. By ranking the most extreme differences between honeybee and fly in terms of domain family occurrences (Supplementary Table 13), the major contributors to phenotypic change become visible (Fig. 7). The most extreme cases in honeybee include expanded families of odorant receptors and the major royal jelly proteins, which are important in caste differentiation, as discussed below.

Apis genes shared with deuterostomes but missing from *Drosophila*. The *D. melanogaster* genome sequence revealed the absence of many genes and pathways that are present in other animals such as mammals and nematodes, and sometimes yeast²⁰. Comparison with the *Anopheles* genome showed that some of these genes are present in other insects and hence must have been lost from the *Drosophila* lineage⁷⁵. Automated comparisons with the OGS suggest 762 proteins that honeybees share with at least one deuterostome but which appear to be missing from *Drosophila*. A similar number are missing from *Anopheles*, but about 300 proteins in *Apis* are missing from both of these dipterans. They were presumably lost early in dipteran evolution or sometime between the divergence of Hymenoptera from the lineage that led to Diptera. These candidate losses of otherwise widespread and conserved animal genes from the *Drosophila* and more broadly fly lineages provide opportunities for many studies. Two of these have been published: the Mahya protein expressed in the mushroom bodies and other brain regions implicated in memory and learning⁷⁹, and pteropsin, a non-visual vertebrate-like opsin expressed deep in honeybee brains and possibly involved in linking the circadian clock to daylight^{80,81}. Other examples, such as telomerase, Dnmt1, Dnmt3 and SID-1 are discussed below.

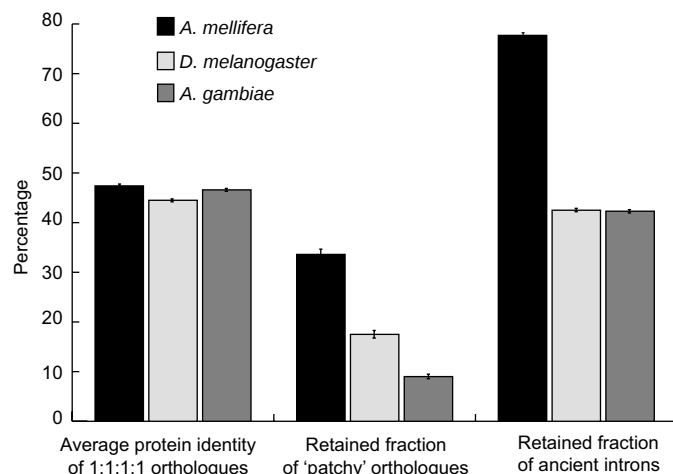


Figure 6 | Comparative evolutionary rates of orthologues. Comparison of single-copy orthologues in honeybee, fly and mosquito versus human in terms of: average protein identity; retained fraction of 'patchy' orthologous groups, as defined in Fig. 5; and fraction of retained ancient introns (those that are found in at least one of the vertebrate orthologues; positional conservation was counted within sliding windows of ± 10 bases to allow for intron sliding). The standard error of the mean is about 0.3% and is shown by the error bars.

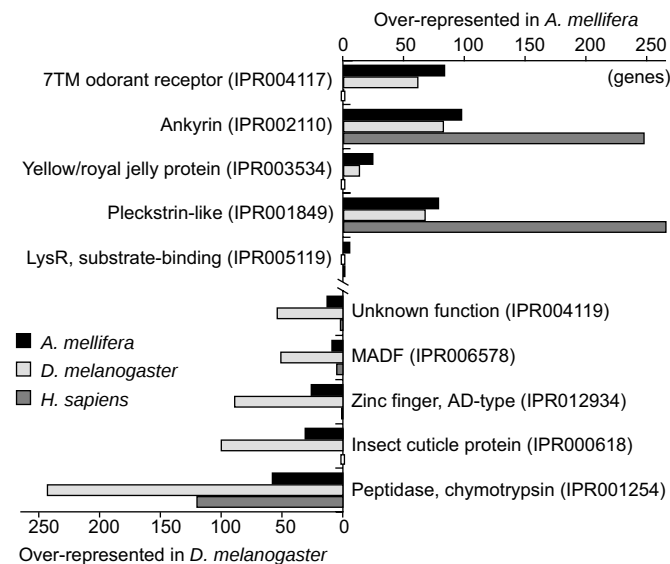


Figure 7 | Protein domains. The top five most prominent expansions and contractions of InterPro-defined protein or domain families in *A. mellifera*, *D. melanogaster* and *H. sapiens*. The families are ordered by the chi-squared test significance of the family size difference with respect to the predicted number of genes, 10,157 and 13,450, respectively. Absent families are marked. 7TM, seven transmembrane.

Functional categories

Using the orthologue set described above, the honeybee genome was compared with other genomes, in particular with the well-annotated, finished *Drosophila* and human genomes to identify protein and domain families for which gene number was different. An important theme in this analysis was relating such differences to the social lifestyle of the honeybee. Table 3 gives some examples of obvious changes whereas Table 4 focuses specifically on neurobiology and behaviour, as discussed below. There also are many examples in which there are subtle changes in the honeybee genome, and some examples including venom components, heat shock proteins, functions for nectar and pollen utilization, and antioxidant systems are presented in the Supplementary Information.

Development

Signalling pathways. A small number of highly conserved cell signalling pathways—Wnt, hedgehog (Hh), transforming growth factor- β (TGF- β), receptor tyrosine kinase (RTK), Notch, Janus kinase (JAK)/signal transducer and activator of transcription (STAT), and nuclear hormone—are responsible for most developmental cell–cell interactions in metazoans⁸². In honeybees, like most metazoans, the genes encoding components of these pathways are conserved; however, the components of some more unusual cell signalling systems, those that specify the early axes in *Drosophila*, are missing from the honeybee genome⁸³. Of the genes that specify terminal embryo fate, *trunk* and *torso* are absent from the honeybee, implying that terminal patterning occurs through a different pathway from that of *Drosophila* and *Tribolium castaneum*. The gene *gurken*, a component

of the *Drosophila* dorso-ventral signalling system, is also missing from the honeybee genome.

This absence of early-acting, axis-specifying genes extends to those that are not involved in cell signalling. Both *bicoid*, an anterior specifying gene, and *oskar* are missing from the honeybee, *Tribolium* and *Bombyx mori* genomes⁸⁴. *Orthodenticle* and *hunchback* genes, which replace *bicoid* in *Tribolium*, are present in the honeybee. Oskar acts as a pole plasm anchor in *Drosophila*⁸⁵, and pole plasm is absent from *Tribolium*⁸⁶, *Bombyx*⁸⁷ and the honeybee⁸⁸.

The absence of some of the earliest acting factors in these pathways is consistent with the hypotheses of refs 89 and 90, which postulate that the initial steps in a developmental cascade are likely to have evolved most recently. It will be important to study how each signalling/patterning system works in the honeybee without these key factors, and whether their activity has been replaced by other, unknown, rapidly evolving genes.

Homeobox genes. Ninety-six homeobox domains were found in 74 genes, either alone or in combination with PAX, POU, LIM and other domains (Supplementary Table 14), similar to *Drosophila*²⁰. The HOX cluster genes are in one cluster on chromosome 16 and the genes are transcribed from the same strand⁸³. This is one of the rare syntenic blocks found in the honeybee and fly genomes.

Sex determination, lack of dosage compensation and male meiosis. The honeybee shows genomic similarities to *Drosophila* for sex determination despite marked differences in this process (Fig. 8)⁸³. Males receive a random half of the mother's genome under a haplodiploid mode of reproduction; there are no gender-specific chromosomes. Sex in the honeybee is determined by the allelic composition of a single locus called complementary sex determiner (*csd*, Fig. 8). Despite the lack of sex chromosomes, honeybees have putative orthologues of some of the genes in *Drosophila*—*run*, *sc* and *dpr*⁹¹—although these are not involved in sex determination in the honeybee and their function is unknown. Most genes downstream in the pathway have orthologues but there is no honeybee *emc* homologue. In the fly cascade, *tra*, which controls somatic sex differentiation, has no orthologue in the honeybee, but the honeybee's initial sex determination signal *csd* is thought to be a functional equivalent of *tra*¹⁹. Orthologues of *dsx* and *ix* are found in the honeybee genome. Honeybee *dsx* is sex-specifically spliced¹⁹, consistent with a conserved sex-determining function in both flies and honeybees. These divergent pathways functionally converge at the *dsx* gene.

Despite the lack of X-specific dosage compensation, the honeybee has orthologues of *mle*, *mof*, *msl-3* and *Trl*, which control dosage compensation in *Drosophila*. It is possible that these genes have additional functions in the honeybee. No potential orthologues of *msl-1*, *msl-2*, *roX1* or *roX2* were identified in the honeybee genome.

Because male honeybees are haploid they lack meiosis. Gene Ontology (GO) analysis identified seven genes in *D. melanogaster* that are involved in male but not female meiosis. Only three of these genes have orthologues in the honeybee (*bol*, *crf*, *topi*). Seven genes specifically involved in female meiosis in *Drosophila* subjected to the same analysis showed only four with orthologues in the honeybee, possibly indicating that several genes involved in the process of meiosis are fast evolving. Most genes of the fly's sex determination pathway are conserved in the honeybee despite the marked differences in

Table 3 | Gene family size differences with possible effects on honeybee lifestyle

Family	Function	Family compared with <i>Drosophila</i>	Possible lifestyle effects
Major royal jelly	Brood feeding	Larger	Brood care; caste development ⁹²
Insulin/insulin-like growth factors	Ageing, fertility, many others	Variable for different subfamilies	Unique reversal of typical lifespan/fertility trade off
Cuticular proteins	Cuticle stability	Smaller	Protected hive environment allows simpler cuticle
Odorant receptors	Olfaction	Larger	Enhanced pheromone communication; odour-based kin recognition; generalist flower feeder
Gustatory receptors	Gustation	Smaller	Brood feeding; mutualistic flower feeder reduces threat of toxic food
Immunity	Infectious disease protection	Smaller	Paradox: high pathogen load due to sociality
Detoxification genes	Defence against xenobiotics	Smaller	Managed environment; specialized lifestyle

Table 4 | Examples of genomic changes with possible impact on brain and behaviour

<i>Drosophila</i> gene(s)	Family	Function	Reference	Novel feature in <i>Apis</i>
Vision				
<i>ninaG</i>	Glucose-methanol-choline oxidoreductases	Involved in biosynthesis of 3-hydroxyretinal	164–166	Two putative <i>ninaG</i> -like genes; different chromophore (11- <i>cis</i> -retinal)
<i>InR</i>	Receptor protein tyrosine kinase	Insulin receptor, photoreceptor-cell (R-cell) axon guidance	167	Gene duplication
<i>norpA</i> <i>Rh1</i> (<i>ninaE</i>) to <i>Rh7</i>	Phospholipase C Rhodopsins	Phototransduction cascade Photoreception	For example, 168 For example, 169	Gene duplication Four visual opsins in honeybee versus seven in fly plus one vertebrate-like non-visual opsin; different chromophore
<i>PNR</i>	Photoreceptor-cell-specific nuclear receptor	Differentiation of the visual system	81, 148	Three PNRs in honeybee versus two in <i>Drosophila</i>
<i>Boss</i>	G-protein-coupled receptor (mGluR-like)	R7 photoreceptor differentiation; ligand of <i>sevenless</i>	170, 171	<i>boss</i> missing but <i>sevenless</i> present; nine photoreceptor cells versus eight in <i>Drosophila</i>
Thermoregulation				
<i>TrpA1</i> , <i>pain</i> (<i>painless</i>), <i>pyx</i> (<i>pyrexia</i>)	TRPA subfamily of transient receptor potential channels	Thermal sensing	172–174	Missing <i>Drosophila</i> <i>TrpA1</i> , but has two extra TRPA channels (GB14005 and GB16385)
Mechanotransduction				
<i>Nach</i> , <i>rpk</i> , <i>ppk</i>	Degenerins/epithelial sodium channels	Mechanotransduction (for example, touch, hearing, nociception)	175	Only 8 degenerins compared with 22 in <i>Drosophila</i>
Learning and memory				
<i>nAcR</i>	Ligand-gated ion channels	Learning and memory processes	103	Eleven nAcR subunits in honeybee instead of ten in <i>Drosophila</i>
<i>Nmdar</i>	Ligand-gated ion channels	Learning and memory processes	176, 177	Three NMDA receptor subtypes in honeybee instead of two in <i>Drosophila</i>
<i>Eaat</i>	Excitatory amino acid transporters	Glutamate uptake	178	Five EAATs in honeybee, two in <i>Drosophila</i>
Circadian rhythms				
<i>tim1</i> (<i>timeless</i>)	Timeless/ Timeout	Circadian clock: component of a feedback loop, light resetting	80	Lost from honeybee
<i>dCry</i> (<i>cryptochrome</i>)	DNA photolyase/ Cryptochrome	Circadian clock: blue-light photoreceptor	80, 179	Lost from honeybee

sexual regulation. It is of interest to understand whether the evolutionary conservation of the genes results from functions other than sex determination that are shared among the honeybee and the fly. This will further support the notion that the early-acting factors of the pathway have been recruited more recently to the sex-determining function and dosage compensation^{89,90}.

Caste, reproduction and ageing

Brood feeding. Royal jelly, produced by exocrine glands in the head of adult worker bees, is an important component of the food used in cooperative brood care, and a key factor in caste differentiation³. The genes encoding the major royal jelly proteins (MRJP) provide one of the best examples of a gene family gaining new functions during the evolution of sociality⁹². The MRJP family in honeybees is encoded by nine genes arranged in a 65-kb tandem array. The MRJP protein family seems to have evolved from a single progenitor gene that encodes a member of the ancient yellow protein family. Five genes encoding yellow-family proteins flank the genomic region containing the genes encoding MRJPs⁹³.

Caste development. Previous studies identified ESTs representing genes differentially regulated during caste differentiation^{94–97}. Improved genome-based annotation of these ESTs and GO analysis of their *Drosophila* orthologues has led to new insights into caste determination.

Genes associated with metabolic regulation are prominent in the EST sets related to caste, corroborating earlier suggestions that changes in metabolism are particularly important in caste determination^{94–97}. For example, queens and workers show different gene expression patterns for oxidoreductases (overexpressed in queen larvae) and hydrolases (overexpressed in worker larvae). In addition, genes overexpressed during worker development are better defined in terms of GO categories than are genes that are overexpressed during queen development (without GO attributes are 0 out of 17 worker overexpressed genes, in contrast to 9 out of 34 genes overexpressed in queens). Even considering the limits in transferring GO terms from *Drosophila* to honeybee, this finding suggests that the evolution of sterility in the worker caste will ultimately be explainable

in terms of known molecular processes. The results also suggest that it may be possible to gain insights from gene expression analyses of caste development in a highly eusocial species into basic questions in socioevolution, namely, what was gained by splitting functions normally performed by a solitary hymenopteran female into two or more castes, and how was this split integrated into post-embryonic differentiation to generate truly alternative phenotypes⁹⁸.

Insulin/insulin-like growth factor signalling. Insulin/insulin-like growth factor signalling pathways are well-conserved integrative pathways regulating ageing, energy metabolism, fertility and other important biological processes. In *Caenorhabditis elegans*, 37 genes encoding putative insulin-like ligands have been identified, whereas 7 have been identified in *D. melanogaster*. However, both *C. elegans* and *D. melanogaster* have a single insulin/IGF-1 receptor orthologue.

There are some notable differences between honeybees and *C. elegans*/*D. melanogaster* for components of insulin/insulin-like growth factor signalling pathways. Honeybees have only two genes encoding insulin-like peptides (ligands) (*AmILP-1* and *AmILP-2*) and two putative insulin/IGF-1 receptors (*AmInR-1* and *AmInR-2*). Honeybees have four orthologues of *C. elegans* *daf-16* (regulator of longevity) compared with one in *Drosophila* (*dFOXO*). These differences in ligand/receptor stoichiometry suggest that honeybees may have evolved a different regulation of this complex pathway. Perhaps these differences also relate to their striking reversal in the traditional relationship between fertility and lifespan. In most organisms, high fertility is achieved at the expense of longevity, whereas in honeybees and other social insects, this relationship is converted into a positive one: queens are both highly fecund and long lived. Recently discovered connections between insulin signalling and the classical insect hormones^{99,100} might provide a link between differential expression of metabolism-related genes in caste development, local patterning in morphogenetic fields and endocrine signals driving metamorphosis. Results from the genome project will increase the effectiveness of the honeybee as a model to examine how the insulin/insulin-like growth factor signalling pathway could have been modified to extend lifespan without negatively affecting reproductive capabilities.

Cuticular and peritrophic membrane proteins. Cuticle consists primarily of chitin and its associated proteins. The most abundant class of cuticular proteins has an extended R&R consensus (pfam00379) that binds to chitin¹⁰¹. The 28 R&R proteins identified in *Apis* are less than one-third of the number of genes for putative cuticular proteins with that domain found in *Drosophila* or *Anopheles*. One possibility is that a less complex cuticle structure allowed by the protected hive environment accounts for the reduced number of genes. This speculation is supported by a comparison of another chitinous structure, the peritrophic membrane of the midgut that has proteins with chitin binding domains (CBD)¹⁰². The *Apis* genome has nine peritrophic membrane proteins with CBDs, a number comparable to those in other insect genomes; some have multiple CBDs.

Telomerase. Organismal ageing is also frequently attributed to a decline in telomerase activity. The presence of the canonical insect telomeric TTAGG repeat at the ends of both kinds of telomeres (distal and proximal) implies that, unlike dipterans, honeybees have telomerase. Using the human telomerase reverse transcriptase (*TERT*) sequence as a query sequence, we identified a candidate gene with 23% amino acid identity to human *TERT*¹⁴. The whole-genome

tiling array (above) and quantitative RT-PCR analysis (M. Corona, H. M. Robertson and G. E. Robinson, personal communication) confirm the expression of telomerase in honeybees. Identification of the honeybee distal telomeres and telomerase will allow study of the possible involvement of telomere length and telomerase activity in the extreme ageing differences of worker, drone and queen honeybees.

Brain and behaviour. Honeybees display a rich behavioural repertoire and have long been recognized as a model system for the study of social interactions. A set of candidate honeybee genes for behaviours representing diverse signalling pathways has already been identified by a number of laboratories¹². Here we explore differences in the diversity of various pathways and gene families previously implicated in brain function and behaviour, and how these might relate to honeybee behaviour (Table 4).

Ion channels, neurotransmitters and other signalling molecules. Similar to the genome of *Drosophila*, the *A. mellifera* genome encodes a conserved set of pore-forming voltage-gated ion channels, but is missing most of the auxiliary subunits found in vertebrates. In contrast to the ~50 two-pore (TWIK) potassium channels present in *C. elegans*, the *A. mellifera* genome encodes only ten. A similar contraction of channel number occurs in the degenerin/amiloride-sensitive sodium channel family, where *A. mellifera* has only 8 genes compared to 24 in *Drosophila*. The classes and numbers of ligand-gated ion channels are largely similar between *Drosophila* and *A. mellifera* (Supplementary Fig. 8), with the interesting exceptions that the *A. mellifera* genome encodes three *N*-methyl-D-aspartate (NMDA) receptor subtypes, instead of two, and five glutamate excitatory transporters, instead of two. The honeybee also has one extra nicotinic acetylcholine receptor (nAChR) subunit¹⁰³. A reduction of genes also holds true for regulatory and catalytic protein kinase A subunits. Moreover, one gene comprises the CREB/CREM family of transcription factors in the honeybee and *Drosophila*, whereas three genes were found in vertebrates¹⁰⁴.

Neuropeptides modulate the behaviour and affect the activity of almost every neuronal circuit. Thirty-six brain peptide genes in *Apis* encode prohormones that are processed into an estimated 200 neuropeptides¹⁰⁵, a number similar to that reported from *Drosophila* and *Anopheles*^{106–108}. However, nine unique neuropeptide genes that do not fit within known neuropeptide families have been found in the honeybee genome. In addition to the neuropeptide genes themselves, 37 neuropeptide and protein hormone G-protein-coupled receptors (GPCRs) have been annotated (48 in *Drosophila*). Furthermore, 19 biogenic amine receptors have been found (21 in *Drosophila*), bringing the total number of honeybee neurohormone GPCRs to 56. The probable ligands for 39 of them have been identified¹⁰⁹.

Apis, like *Drosophila*, lacks brain-derived neurotrophic factor (BDNF) signalling machinery, which is well conserved in vertebrates. However, the honeybee has all the components of the agrin synapse formation pathway, several of which, including agrin itself, are missing in flies. The core synaptic vesicle trafficking machinery is largely conserved between the currently sequenced invertebrate genomes, although the honeybee synaptotagmin family is more similar to mice and humans than *Drosophila* or *Anopheles*.

Chemoreceptors. In the honeybee genome there is a remarkable expansion of the insect odorant receptor family relative to *D. melanogaster* (62 odorant receptors from 60 genes)^{110–112} and *A. gambiae* (79 odorant receptors)¹¹³. A total of 170 odorant receptor genes were manually annotated, of which 7 are pseudogenes¹¹⁴ (this is twice the number of seven transmembrane odorant receptors in Fig. 7 because the automated annotations concatenated many tandemly arrayed genes). These constitute five honeybee-specific subfamily lineages in an insect odorant receptor family tree, and one of these lineages is hugely expanded with 157 genes encoding 15–99% amino acid identity. One-hundred and forty-two of these odorant receptor genes are in 14 tandem arrays of two or more genes distributed throughout the genome, including one of 60 genes, reflecting likely expansion by

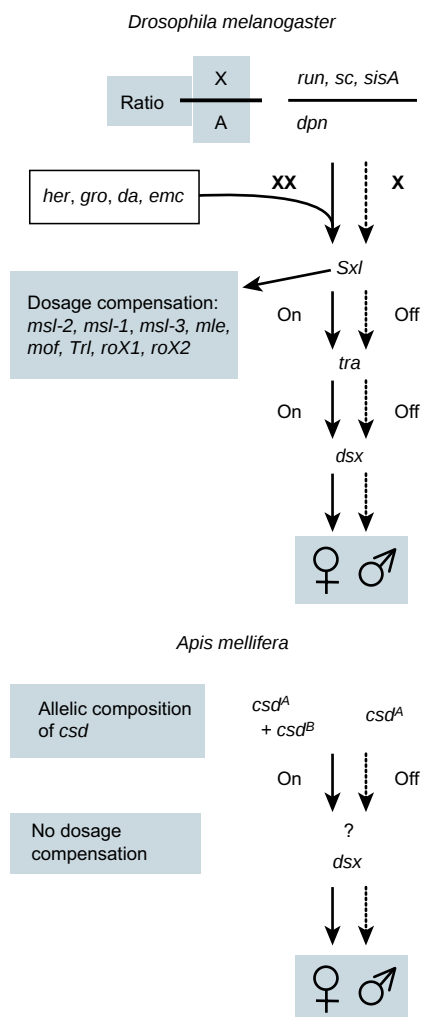


Figure 8 | Sex-determining pathways of *Drosophila* and *Apis*. Sex in the honeybee is determined by the allelic composition of a single gene, the complementary sex determiner (*csd*)¹⁹. Eggs develop into males when *csd* is hemizygous (haploid) or homozygous, or females when *csd* is heterozygous. Honeybees lack sex chromosomes and X-specific dosage compensation. Sex-specific information is transferred in both species from diverged initial signals to the final gene, *dsx*, via switch genes that are active (on) in the females, but inactive (off) in the males. Most *Drosophila* pathway genes are present in the honeybee genome despite the marked differences (see text).

unequal crossing over. This huge odorant receptor family expansion presumably mediates the honeybee's remarkable range of odorant capabilities, including perception of several pheromone blends, subtle kin recognition signals, and diverse floral odours. It is notable that this large number of odorant receptor genes corresponds well with the estimated number of 160–170 glomeruli in the honeybee antennal lobe¹¹⁵, consistent with the central model of insect olfaction, which holds that each olfactory receptor neuron expresses one¹¹² or at most two¹¹⁶ odorant receptors. All olfactory receptor neurons expressing the same odorant receptors then converge on a single glomerulus per antennal lobe, allowing for an odour map of patterns of olfactory receptor neuron stimulation¹¹⁵.

The odorant receptor family is but one highly expanded lineage of a superfamily of several distantly related lineages of chemoreceptors¹¹¹, most of which are implicated in gustatory function and therefore are called gustatory receptors¹¹⁷, although some might have olfactory functions¹¹⁶. In contrast to the *D. melanogaster* repertoire of 68 gustatory receptors encoded by 60 genes¹¹¹ and the *A. gambiae* repertoire of 76 gustatory receptors encoded by 52 genes¹¹³, the honeybee gustatory receptor repertoire consists of just 10 unclustered gustatory receptors representing 7 divergent lineages¹¹⁴. The honeybee might have lost one or two ancient gustatory receptor lineages, but the difference is largely the result of differential expansion of subfamilies in the two flies and lack of any gustatory receptor subfamily expansion in the honeybee. The limited gustatory receptor repertoire perhaps reflects the provisioning of larval bees by adults and their reduced need to avoid toxic chemicals in their food—the relationship between honeybee and flower is mutualistic. Honeybees also antennate each other and other objects, and are perhaps using some odorant receptors as gustatory receptors.

Odorant binding proteins are a third major component of the insect chemosensory system, with the potential to influence chemosensation¹¹⁸. The honeybee genome encodes 21 odorant binding proteins, less than half the number of either dipteran, with 51 in *Drosophila* and 70 in *Anopheles*¹¹⁹. If odorant binding proteins function in a combinatorial fashion with the receptors, this relatively limited repertoire might restrict the range of honeybee olfaction; however, the roles of odorant binding proteins beyond transport of hydrophobic molecules in insect chemosensation remain unclear and many are not expressed in chemosensory organs.

Circadian rhythms. The honeybee orthologues of the clock proteins cryptochrome (Cry), timeless (Tim), clock (Clk) and cycle (Cyc), which are mostly associated with the 'negative feedback' loop of the clock, are more similar to mammals than *Drosophila*⁸⁰. For example, the honeybee genome does not contain orthologues to *dCry* (*Drosophila*-type Cry) and *timeless1* (*tim1*), genes that are essential for clock function in the central pacemaker of *Drosophila*, but does have orthologues to genes encoding the mammalian-type paralogues mCry (mammalian-type Cry) and timeout (Tim2), which are thought to have different clock function¹²⁰. The temporal brain expression pattern of *AmCyc*, *AmClk*, *AmCry* and *AmTim2* is more similar to mammals than to *Drosophila*, suggesting that they behave like their mammalian orthologues. Additional analysis suggests greater *Apis*–*Drosophila* similarities in the 'positive feedback loop' of the clock. The honeybee genome encodes highly conserved orthologues to Vri1 (Vri) and PAR domain protein 1 (Pdp1), two basic zipper transcription factors that are implicated in the regulation of *Clk* expression in *Drosophila*, with the highest similarity (>94% identity, 100% similarity) in the DNA binding domain. On the other hand, the honeybee genome does not encode true orthologues to the orphan nuclear receptors Rev-Erb (α and β ; nuclear receptor subfamily 1, group D members 1 and 2; NR1D1 and NR1D2, respectively) and ROR (α , β , γ ; nuclear receptor subfamily 1, group F, members 1–3; NR1F1–3, respectively), which are thought to orchestrate the expression of Bmal1 (the vertebrate orthologue of cycle) in the mammalian clock¹²¹. Although the honeybee genome encodes related nuclear receptors (GB11364-PA, related to Ecdysone-induced

protein 75B, a NR1D3 protein, and GB10650-PA, similar to Hormone receptor like 45, a NR1F4 protein), we found PAR family consensus elements (putative binding sites for Vri and Pdp1) but no RORE response elements (putative binding sites for ROR and Rev-Erb proteins) in the 3 kb upstream of the start codon of either *AmClk* or *AmCyc*. Phylogenetic analyses show that the basal animal lineage had both the mammalian and *Drosophila* types of *Cry* and *Tim*⁸⁰ (Supplementary Fig. 9). Thus, *Drosophila* diverged by specializing on using one set of orthologues and by losing mCry; honeybees converged with mammals by losing these orthologues and specializing on the other set. These analyses of the honeybee genome sequence uncover previously unknown diversity in animal clocks, challenge the distinction commonly made between the clocks of insects and vertebrates, and raise critical questions concerning the evolution and function of clock genes.

Immunity and disease resistance. Honeybees live in highly crowded nests, providing favourable demographic conditions for infectious diseases. Honeybee pathogens are well known¹²² and include viral, bacterial, fungal and protist pathogens, along with other parasites¹²³. Protection from infectious disease includes social defences such as grooming and other hygienic behaviours, individual chambers for raising young, and a workforce that defends the nest against many potential vectors of disease. Individual honeybees are also defended by morphological barriers and immune defences.

Curiously, given the predicted disease pressures in honeybee colonies, the honeybee genome encodes fewer proteins implicated in insect immune pathways when compared to other insect genomes. Although the Toll, Imd and JAK/STAT pathways seem to be intact, paralogue counts for gene families implicated in these pathways are reduced by two-thirds¹²⁴. This reduction spans every step in the immune response from pathogen recognition to immune effectors, and implies a reduced flexibility in the abilities of honeybees to recognize and resist pathogens. The results suggest that honeybees use novel immune pathways, are poorly defended against pathogens at the individual level, and/or have immune systems that are narrowly focused on a relatively small group of coevolved pathogens.

Gene-expression and RNAi studies have begun to elucidate roles for several candidate immune genes in response to microbial infection^{125,126}. Coupling these studies with efforts to understand pathogen gene expression^{127,128} and invasion mechanisms¹²⁹, and with the long-standing search for the impacts of genotypic and environmental variance on disease resistance^{73,75}, will cement the honeybee as an essential model for the study of immunity and disease in social insects.

Anti-xenobiotic defence mechanisms. Honeybees are vulnerable to insecticides and have suffered major population losses in some regions of the world. Contact pesticides affect the worker bees whereas residual pesticides accumulate in lipophilic substances such as wax or pollen lipids¹³⁰ and impact on the developing brood and queen fecundity. Wax acts as a pesticide sink¹³¹ and pesticide residues incorporated into wax may migrate to honey¹³². Sub-lethal effects of pesticides affect honeybee initial learning and conditioned odour responses, traits directly linked to foraging.

It seems that the size of the major detoxifying gene families is smaller in the honeybee, making the species unusually sensitive to certain pesticides^{133–136}. Compared with *Anopheles* and *Drosophila*¹³⁷, the honeybee has 30–50% fewer genes encoding the carboxylesterase (CCE), cytochrome P450 (P450) and glutathione S-transferase (GST) enzymes that are principally responsible for the metabolism of pesticides and in which the great majority of metabolic resistance mutations have been found in other species of invertebrates¹³⁸. The greatest difference is seen in the GSTs, the family most strongly associated with detoxifying functions. Two clades of GSTs containing all known insecticide-resistance-related GSTs in other insects, and comprising over 20 members in *D. melanogaster* and *A. gambiae*, consist of a single member in the honeybee. Similarly, the honeybee genome has just half the number of P450 genes and contains less than

20% of genes of the CYP4 clade (21 in *Drosophila*, 4 in the honeybee), which is strongly associated with pyrethroid resistance in other species.

Conversely, there is usually the same number of genes, and in some cases significantly more, in clades within the CCE, P450 and GST families that are not involved in detoxification. The marked variation in relative abundance argues strongly against explanations for changes in gene number in honeybee based on genome-wide factors like its haplodiploid genetic system and the associated exposure to selection in the haploid caste.

Gene regulation

DNA methylation. DNA methylation systems are well characterized in vertebrates¹³⁹. By contrast, methylation in *D. melanogaster* and other invertebrates remains controversial¹⁴⁰. Evidence for DNA methylation has been demonstrated in several different orders of insects, but these results are interpreted cautiously because no catalytically active deoxycytosine methyltransferase has yet been identified and characterized in any invertebrate. The honeybee genome contains genes that encode orthologues of all vertebrate proteins required for DNA methylation¹⁴⁰. In addition to Dnmt2 (also found in Diptera), three CpG-specific DNMT family genes were identified: two Dnmt1 genes and one Dnmt3a/b gene. A single putative methyl-DNA-binding-domain-containing gene with two splicing variants was also identified. The DNMT genes are expressed and are active *in vitro*. Moreover, the honeybee methylation system is functional *in vivo*, as shown by genomic 5-methyl deoxycytosine in honeybee DNA and several specific CpG methylated endogenous gene sequences¹⁴⁰.

RNA interference. The honeybee genome harbours only a single homologue of *sid-1*, a gene essential for systemic RNAi¹⁴¹, whereas multiple copies of this gene occur in the moth and draft *Tribolium castaneum* genomes⁸⁴. The honeybee SID-1 homologue clusters with one of three beetle and two of three moth SID homologues (Fig. 9), as well as two vertebrate paralogues, one of which was lost in fish but preserved in birds and mammals.

This disparity in SID-1-like gene number may relate to honeybee RNAi efficacy. RNAi in honeybees was first demonstrated by embryo injection of double-stranded RNAs¹⁴², and later experiments showed that, as in other invertebrates, injection of double-stranded RNA can reduce target gene expression away from the site of injection¹⁴³. However, the molecular details of honeybee gene silencing have not

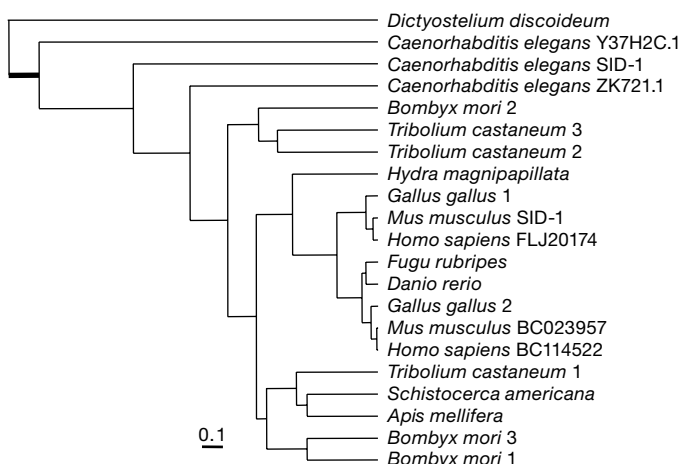


Figure 9 | Phylogenetic analysis of the SID-1 proteins of insects and diverse other organisms. The maximum likelihood phylogram (Phylip) is based on the alignments of the conserved carboxy-terminal transmembrane domain¹⁴¹ of SID-1 proteins from insects, vertebrates, nematodes and other eukaryotes. Sequences were aligned with Clustalw. The sequences are identified by species name and, where multiple genes exist in a genome, by a further identifier (see Supplementary Information). A long basal branch from *Dictyostelium* to the other sequences is truncated (thick bar).

been clearly elucidated¹⁴⁴, and little is known about the role of SID-1 proteins in insect systemic RNAi.

The honeybee genome also encodes many other proteins similar to the core RNAi machinery of other species, including two Dicer (Dcr) enzymes, a Drosha homologue and other RISC components, like argonaute 1 and 2, as well as a full suite of double-stranded RNA binding proteins—R2D2, Pasha and Loquacious—all of which have yet to be confirmed as components of the RNAi pathway in honeybees.

Nuclear receptor transcription factors. The set of 22 nuclear receptors encoded by the *Apis* genome is a nearly perfect match to the *Drosophila* set, a strong reminder of the centrality of these transcription factors in the regulation of insect embryonic development and metamorphosis¹⁴⁵. The single novelty uncovered in *Apis* nuclear receptors is subtle: the presence in *Apis* of a third gene homologous to vertebrate photoreceptor-cell-specific nuclear receptor (PNR), represented in *Drosophila* by DHR51 and DHR83 (ref. 146). If, as predicted on the basis of studies of human PNR mutations, *Apis* PNR-like is involved in the differentiation of the visual system^{147,148}, this additional gene may reflect a key aspect of the honeybee's behavioural ecology: reliance on vision for navigation. This finding is consistent with an apparent overall increase (relative to *Drosophila*) in the number of genes putatively involved in photoreception⁸¹.

cis-Regulation of behavioural development. Adult worker honeybees typically shift from working in the hive to foraging for nectar and pollen outside the hive when they are 2–3 weeks of age, but the age at which this transition is made depends on the needs of the colony, which are communicated among honeybees via pheromones¹². To begin to study the *cis*-regulatory code associated with this form of social regulation, the genome sequence was scanned¹⁴⁹ for regulatory motifs in the promoter regions of genes expressed in the brain that are related to socially regulated behaviour development and identified via microarray analysis^{150,151}.

Results of statistical analyses (see Supplementary Information) indicate that the transcriptional regulatory pathway for social behaviour in honeybees shares several commonalities with the pathway for development in *Drosophila*. Gene expression was significantly associated with binding sites for the transcription factors Hair, GAGA, Adf1, CF1, Snail and Dri. Promoter sequence patterns predicted brain expression for as many as 71% of certain types of behaviour-related genes, even though less than 15% of all transcription factors known from *Drosophila* were studied. Social regulation of gene expression is a potent influence on behaviour in animals and humans¹², and these results will help to elucidate underlying mechanisms.

Bee phylogeny and population genetics

Information from the honeybee genome is providing new insights into the origins of honeybees and their spread throughout the world. Ordinal relationships of the higher insects have been re-examined using concatenated sequences from 185 (ref. 152) and 1,150 (ref. 76) protein-coding genes. These new phylogenies support the hypothesis that the order Hymenoptera (ants, bees and wasps) is derived from an early branch in the holometabolous insects, which challenges the current view that the lineage leading to Hymenoptera diverged after the evolution of beetles (Coleoptera) (Fig. 1). Phylogenetic analysis of all bee families using sequences derived from the Honey Bee Genome Project supports a hypothesis that the Apidae branched off earlier in the phylogeny of bees than previously recognized¹⁵³. This, combined with the fact that the oldest fossil bee (*Cretotrigona prisca*) is a corbiculate apid, supports the view that honeybees and their relatives are the oldest lineage of eusocial bees. Other eusocial bee lineages (Halictidae and Allodapini) evolved eusociality later than honeybees¹⁵⁴.

The honeybee genome sequence was used as the basis for the development of a large set of SNPs. These have been used to generate a view of the relationships among *A. mellifera* subspecies of

unprecedented detail. In the following sections we summarize the SNP resource, present an initial view of the relationships among *A. mellifera* subspecies, and summarize key findings³⁰ on historical patterns of migration, differentiation and introgression of honeybees in both the Old and New World to help define the origins of domesticated honeybees and their resultant worldwide diaspora. The SNP set will also be essential for marker-assisted and positional cloning of genes that underlie important behavioural, social and economically relevant traits.

SNP resources. Candidate SNPs were identified from two sources: (1) alignment of 2,483 genome traces from Africanized honeybees to the assembled genome sequence (European-derived); and (2) alignment of ~75,000 ESTs derived from genetically variable European honeybees. Using a bayesian method¹⁵⁵, 3,594 and 1,950 base-substitution SNPs from each source (respectively) showed a high probability ($P > 0.99$) of being true polymorphisms. An initial honeybee SNP panel was generated from 1,536 putative SNPs, of which 1,136 were validated (see Supplementary Information). These SNPs were genotyped in 328 *A. mellifera* (from native and introduced populations) and 13 individuals from three related species: *Apis cerana*, *Apis dorsata* and *Apis florea*. Details of these methods are presented in Supplementary Information. Results are summarized in the following section and described in more detail elsewhere³⁰.

Biogeography and phylogeny of native and Africanized honeybees. Analyses of SNP genotypes in ten subspecies (Fig. 10) revealed four non-overlapping clusters of *A. mellifera* subspecies analogous with four evolutionary lineages (called M, C, O and A) defined by morphometric characters^{28,156} and consistent with mitochondrial DNA analysis¹⁵⁷ (where C and O lineages, as defined here and in ref. 28, correspond to the 'C' mitochondrial type). Surprisingly, within *A. mellifera*, north and west European honeybees (M) were more similar to African (A) than to the geographically proximal east European honeybees (C) (Fig. 10). Phylogenetic analysis using outgroup genotypes from *A. cerana*, *A. dorsata* and *A. florea* suggested that extant *A. mellifera* subspecies originated in Africa³⁰, consistent with speculation of an origin in tropical or subtropical Africa by ref. 2 but contrary to current hypotheses of an origin for *A. mellifera* in western Asia¹⁵⁶ (based primarily on the occurrence of approximately ten allopatric species in eastern Asia). Taken together, these data support a hypothesis involving an African origin for *A. mellifera* and at least two separate migrations into Eurasia: a migration into Europe via the Iberian Peninsula, which has since expanded into central Europe

and Russia (M group), and one (or more) migrations into Asia and east Europe south of the Alps (O and C groups, respectively).

Africanization in the New World has involved the near-replacement of the 'European' honeybee by descendants of *A. m. scutellata*. Analysis of SNP genotypes in New World bees³⁰ revealed several key findings. First, pre-existing 'European' populations showed evidence of extensive admixture between C, M and O groups (consistent with known introductions from at least nine subspecies¹⁵⁸). Second, although African alleles were dominant in populations after Africanization, all Africanized individuals showed evidence of introgression between *A. m. scutellata* and pre-existing populations. Third, replacement of European- by African-derived alleles was origin-dependent: pre-existing C group alleles were replaced but M group alleles were not. The explanation for the latter result is unclear, but may result from the close relationship between A and M (compared to A and C; Fig. 10) or from historical and local patterns of introductions.

Conclusions

After Mendel completed his work with peas, he turned to experimenting with honeybees, to extend his work to animals (<http://www.zephyrus.co.uk/gregormendel.html>)¹⁵⁹. He produced a hybrid strain (so vicious they were destroyed) but failed to reproduce the clear picture of heredity because of difficulties in controlling the mating behaviour of queens¹⁵⁹. The honeybee genome similarly proved a challenge for sequencing with biases in clone libraries. With special remedies we produced a draft sequence that is high quality by a number of metrics. The overall statistics for completeness (genome size, coverage of markers, ESTs and cDNAs) indicate that over 96% of these important elements are in the sequence. Similarly, assessments of quality indicated few misassemblies, mainly associated with repeated sequences, as expected for a draft sequence. We expect that regions remain with lower sequencing redundancy but these regions are not missing altogether and their sequences are included in homology searches and can be recognized. Other expected limitations are accurate placement of repeated sequences and heterochromatic regions, which are problematic even in sequences taken to the highest finished quality.

In addition to providing sequence, a genome project also produces a gene list, generated here by novel merging of five efforts⁶⁹. The OGS consisted of a little more than 10,000 genes, lower than other insects (*D. melanogaster*, 13,600 genes²⁰; *A. gambiae*, 14,000 genes²³; *B. mori*, 18,500 genes^{26,160}). Genome tiling array and manual annotation efforts increase the OGS by a few per cent, but it remains smaller by 15% or more compared with other sequenced insects. Because the sequence does not appear to be missing such an amount of the genome, we believe that the gene count is underestimated. Consistent with this are results from whole-genome tiling array experiments that detected signals in thousands of regions currently described as intergenic.

We suspect two reasons for the current low gene number. First, limited EST and cDNA data for *Apis* reduced gene predictions. Second, the large evolutionary distance of *Apis* from other sequenced genomes restricted use of orthology in predicting genes and may have introduced a bias in the OGS. When genes that are known in other organisms were not found in the OGS or the OAIGS, orthologues themselves were used to search the honeybee genome, and genes found were added to the OGS. But shorter genes, rapidly diverging genes, and other special cases may not be readily detected by this approach.

Members of honeybee gene families were less likely to be missed because other family members showed sufficient sequence similarity to be useful in searches. Supporting this is the observation of a number of expanded gene families; for example, the major royal jelly protein and odorant receptor families. We predict that the gene count for the honeybee will increase in the future as more data and analyses are applied.

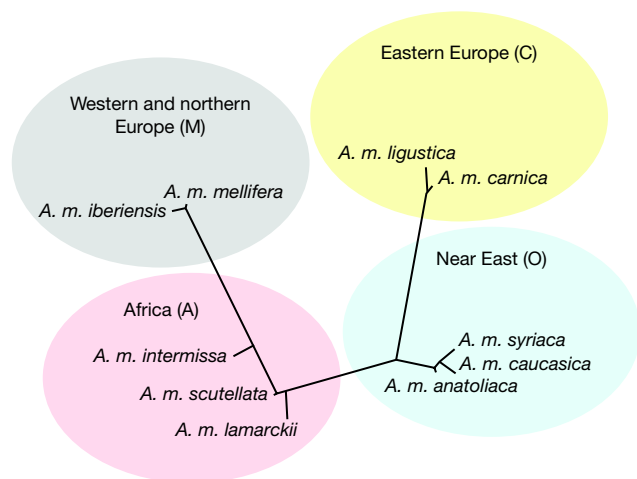


Figure 10 | Population genetic structure of honeybees collected from native ranges in Europe, Africa and the Near East. Neighbour-joining tree using Nei genetic distance¹⁸¹. Ten geographical subspecies ($N = 9\text{--}21$ individuals each) can be partitioned into four regional groups. Branches separating regional groups are supported by 100% bootstrap.

The honeybee genome's high A+T content, absence of transposons and slower rate of evolution, will be better understood as more insect genomes are sequenced¹⁶¹. The paradoxes of high CpG content despite the presence of cytosine methylases, fewer genes for innate immunity despite the high pathogen and parasite loads associated with social life, or the presence of *Drosophila* orthologues for many genes in the sex determination pathway despite the honeybee's lack of sex chromosomes, promise to establish the honeybee as a new model for several fundamental processes of life.

A number of new resources have been produced to enhance biological discovery. The discovery of genes related to RNAi (for example, a SID-1-like protein) should enhance the use of this technique in the honeybee. It is difficult to select and propagate mutants in the haplodiploid, polyandrous, open-air-mating honeybee, making it difficult to apply traditional forward genetic tools. RNAi can now be used for clarifying gene function and altering protein expression for beneficial effects on honeybee behaviour or physiology. Equally important, the genome facilitates characterization of core molecular features of RNAi, potentially clarifying inconsistencies observed in silencing some honeybee gene targets. The discovery of novel miRNAs and *cis*-regulatory elements associated with social behaviour provides a foundation to begin to understand social regulation of gene expression. Another new resource is an extensive SNP set, which already has generated new insights into honeybee phylogeography and invasion biology, and will prove invaluable for positional cloning of genes in quantitative trait loci for a variety of traits such as defensive behaviour or foraging behaviour¹⁶².

How will the honeybee genome sequence enable a mechanistic understanding of social organization, communication and the ability to shape the local environment? The evolution of sociality requires changes to every system in the organism, not only to invent new functions but also to tune old ones to new purposes. We expected to find a rich set of genetic features underpinning honeybee sociality, and here again, the genome project has not disappointed.

One intriguing trend is a smaller size of some gene families relative to the other sequenced insect genomes, possibly reflecting a selective elimination of genes whose functions have become superfluous in the now highly specialized life history and self-managed environment of the honeybee. Larger gene family sizes are, however, also observed. New genes are not created *de novo*, but result from duplication and diversification. The initial analysis of the honeybee genome presented here shows 60 such duplications that are not present in other genomes. These, and others like them awaiting discovery, are candidates for honeybee-specific functions. The major royal jelly proteins provide a good example of protein family expansion and social evolution⁹³.

However, achieving a comprehensive understanding of social life in molecular terms will require extensive analyses of the honeybee as well as other social and non-social species. A genome might be a blueprint for some aspects of biology, but most mysteries of sociality appear to be encoded subtly in the genome, at least based on our study of honeybee and *Drosophila*, as well as recent analyses of human and chimpanzee¹⁶³. Although much remains to be done, with the genome in hand, and the associated methodologies it enables, prospects are bright for elucidating the molecular and genetic bases of many complex traits associated with honeybee sociality.

METHODS

Detailed methods are described in Supplementary Information. Sources for resources generated by this project are listed here.

Genome assemblies. Genome assemblies are available from the BCM-HGSC ftp site under the directory (<ftp://ftp.hgsc.bcm.tmc.edu/pub/data/Amellifera/fasta/>) (see Supplementary Table 17). The files available for each assembly differ, but in general there are directories for contigs (the sets of contigs with fasta quality and gap files for each linkage group), for linearized scaffolds (sequences for each linkage group where the gaps between contigs have been filled with Ns), for bin0 (non-overlapping reads) and repeat reads. The file descriptions and assembly statistics for each version are described in a file named README.txt.

The version 4.0 assembly directory also contains small contigs (less than 1 kb) that were omitted from the assembly as well as haplotype contigs (overlapping contigs identified as representing the second haplotype in the sequenced DNA). The individual accessions at the NCBI are: version 4.0 (scaffolds CM000054–CM000069, CH876891–CH878241); version 3.0 (contigs AADG05000001–AADG05018946); version 2.0 (contigs AADG04000001–AADG04016028; scaffolds CM000054–CM000069, CH402995–CH404444); version 1.2 assembly (contigs AADG03000001–AADG03022771; scaffolds CM000054–CM000069, CH236967–CH239577); version 1.0 (contigs AADG02000001–AADG02030074); version 1.0 (contigs AADG01000001–AADG01015795). The version 2.0 assembly is displayed in the NCBI Map Viewer.

Genome browsers. Genome browsers are available for viewing the genome assemblies at BeeBase (http://racexr00.tamu.edu/bee_resources.html), NCBI (<http://www.ncbi.nlm.nih.gov>) and UCSC (<http://genome.ucsc.edu>). The assembly versions are listed in Table 1 and Supplementary Table 1. The feature annotations that are available differ from site to site.

Chromosome superscaffolds. Chromosome superscaffolds for chromosomes 13, 14, 15 and 16 are available at BeeBase (http://racexr00.tamu.edu/bee_resources.html).

SNPs. SNPs identified from *A. m. scutellata* whole-genome shotgun (WGS) reads and from the two haplotypes within the *A. m. mellifera* assembly are available from dbSNP at NCBI (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=SNP&cmd=Limits>) and from the BCM-HGSC ftp site (<ftp://ftp.hgsc.bcm.tmc.edu/pub/data/Amellifera/snp>). SNPs identified from EST sequences are available from UIUC (http://titan.biotec.uiuc.edu/bee/downloads/bee_downloads.html).

Tiling array data and EST sequences. Tiling array data are available for browsing and download from Systemix (<http://www.systemix.org>). EST sequences are available from NCBI, DDBJ and EMBL under accessions DB728206–DB781564.

Gene predictions. All of the individual gene sets in the OAIGS, OGS, the community annotated set and BeeBase manually curated set are available for download from the BeeBase downloads page (<http://racexr00.tamu.edu/downloadFASTA.html>) as either protein or CDS sequences.

Received 13 July; accepted 19 September 2006.

- Wilson, E. O. & Holldobler, B. Eusociality: origin and consequences. *Proc. Natl Acad. Sci. USA* **102**, 13367–13371 (2005).
- Wilson, E. O. *The Insect Societies* (Harvard Univ. Press, Cambridge, 1971).
- Winston, M. L. *The Biology of the Honey Bee* (Harvard Univ. Press, Cambridge, 1987).
- Evans, J. D. & Wheeler, D. E. Gene expression and the evolution of insect polyphenisms. *Bioessays* **23**, 62–68 (2001).
- Page, R. E. Jr & Peng, C. Y. Aging and development in social insects with emphasis on the honey bee, *Apis mellifera* L. *Exp. Gerontol.* **36**, 695–711 (2001).
- Witthöft, W. Absolute Anzahl und Verteilung der Zellen im Hirn der Honigbiene. *Z. Morphol. Tiere* **61**, 160–184 (1967).
- Menzel, R. Searching for the memory trace in a mini-brain, the honeybee. *Learn. Mem.* **8**, 53–62 (2001).
- von Frisch, K. *Dance Language and Orientation of the Honey Bee* (Harvard Univ. Press, Cambridge, 1967).
- Giurfa, M., Zhang, S., Jenett, A., Menzel, R. & Srinivasan, M. V. The concepts of 'sameness' and 'difference' in an insect. *Nature* **410**, 930–933 (2001).
- Sheppard, W. S., Rinderer, T. E., Garnery, L. & Shimanuki, H. Analysis of Africanized honey bee mitochondrial DNA reveals further diversity of origin. *Genet. Mol. Biol.* **22**, 73–75 (1999).
- Smith, D. R. & Brown, W. M. Polymorphisms in mitochondrial DNA of European and Africanized honeybees (*Apis mellifera*). *Experientia* **44**, 257–260 (1988).
- Robinson, G. E., Grozinger, C. M. & Whitfield, C. W. Sociogenomics: social life in molecular terms. *Nature Rev. Genet.* **6**, 257–270 (2005).
- Honey Bee Genome Sequencing Consortium. Proposal for the sequencing of a new target genome: White paper for a honey bee genome project (http://www.genome.gov/Pages/Research/Sequencing/SeqProposals/HoneyBee_Genome.pdf) (2002).
- Morse, R. A. & Calderone, N. W. The value of honey bee pollination in the United States. *Bee Culture* **128**, 1–15 (2000).
- Hamilton, W. D. The genetical evolution of social behaviour, I, II. *J. Theor. Biol.* **7**, 1–52 (1964).
- Hamilton, W. D. Altruism and related phenomena, mainly in social insects. *Annu. Rev. Ecol. Syst.* **3**, 193–232 (1972).
- Crozier, R. H. & Pamilo, P. *Evolution of Social Insect Colonies: Sex Allocation and Kin Selection* (Oxford Univ. Press, Oxford, 1996).
- Foster, K. R., Wenseleers, T. & Ratnieks, F. L. W. Kin selection is the key to altruism. *Trends Ecol. Evol.* **21**, 57–60 (2006).
- Beye, M., Hasselmann, M., Fondrk, M. K., Page, R. E. & Omholt, S. W. The gene *csd* is the primary signal for sexual development in the honeybee and encodes an SR-type protein. *Cell* **114**, 419–429 (2003).
- Adams, M. D. et al. The genome sequence of *Drosophila melanogaster*. *Science* **287**, 2185–2195 (2000).

21. Misra, S. *et al.* Annotation of the *Drosophila melanogaster* euchromatic genome: a systematic review. *Genome Biol.* **3**, RESEARCH0083.1–RESEARCH0083.22 (2002).
22. Richards, S. *et al.* Comparative genome sequencing of *Drosophila pseudoobscura*: chromosomal, gene, and cis-element evolution. *Genome Res.* **15**, 1–18 (2005).
23. Holt, R. A. *et al.* The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science* **298**, 129–149 (2002).
24. Mongin, E., Louis, C., Holt, R. A., Birney, E. & Collins, F. H. The *Anopheles gambiae* genome: an update. *Trends Parasitol.* **20**, 49–52 (2004).
25. Mita, K. *et al.* The genome sequence of silkworm, *Bombyx mori*. *DNA Res.* **11**, 27–35 (2004).
26. Xia, Q. *et al.* A draft sequence for the genome of the domesticated silkworm (*Bombyx mori*). *Science* **306**, 1937–1940 (2004).
27. Grimaldi, D. & Engel, M. S. *Evolution of the Insects* (Cambridge Univ. Press, Cambridge, 2005).
28. Ruttner, F. *Biogeography and Taxonomy of Honeybees* (Springer, Berlin, 1988).
29. Garnery, L., Cornuet, J. M. & Solignac, M. Evolutionary history of the honey bee *Apis mellifera* inferred from mitochondrial DNA analysis. *Mol. Ecol.* **1**, 145–154 (1992).
30. Whitfield, C. W. *et al.* Thrice out of Africa: ancient and recent expansions of the honey bee, *Apis mellifera*. *Science* (in the press).
31. Havlak, P. *et al.* The Atlas genome assembly system. *Genome Res.* **14**, 721–732 (2004).
32. Solignac, M. *et al.* A microsatellite-based linkage map of the honeybee, *Apis mellifera* L. *Genetics* **167**, 253–262 (2004).
33. Solignac, M. *et al.* The genome of *Apis mellifera*: dialog between linkage mapping and sequence assembly. *Genome Biol.* (submitted).
34. Jordan, R. A. & Brosemer, R. W. Characterization of DNA from three bee species. *J. Insect Physiol.* **20**, 2513–2520 (1974).
35. Beye, M. & Raeder, U. Rapid DNA preparation from bees and %GC fractionation. *Biotechniques* **14**, 372–374 (1993).
36. Human Genome Sequencing Center at Baylor College of Medicine. Honey Bee Genome Project (<http://www.hgsc.bcm.tmc.edu/projects/honeybee>) (2006).
37. Beye, M. *et al.* Exceptionally high levels of recombination across the honey bee genome. *Genome Res.* (in the press).
38. Hunt, G. J. & Page, R. E. Jr. Linkage map of the honey bee, *Apis mellifera*, based on RAPD markers. *Genetics* **139**, 1371–1382 (1995).
39. Beye, M. & Moritz, R. F. Characterization of honeybee (*Apis mellifera* L.) chromosomes using repetitive DNA probes and fluorescence *in situ* hybridization. *J. Hered.* **86**, 145–150 (1995).
40. Cohen, N., Dagan, T., Stone, L. & Graur, D. GC composition of the human genome: in search of isochores. *Mol. Biol. Evol.* **22**, 1260–1272 (2005).
41. Duret, L., Mouchiroud, D. & Gautier, C. Statistical analysis of vertebrate sequences reveals that long genes are scarce in GC-rich isochores. *J. Mol. Evol.* **40**, 308–317 (1995).
42. Matsuzaki, M. *et al.* Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. *Nature* **428**, 653–657 (2004).
43. Coulondre, C., Miller, J. H., Farabaugh, P. J. & Gilbert, W. Molecular basis of base substitution hotspots in *Escherichia coli*. *Nature* **274**, 775–780 (1978).
44. Robertson, H. M. & Gordon, K. H. J. Canonical TTAGG repeat telomeres and telomerase in the honey bee, *Apis mellifera*. *Genome Res.* (in the press).
45. Sahara, K., Marec, F. & Traut, W. TTAGG telomeric repeats in chromosomes of some insects and other arthropods. *Chromosome Res.* **7**, 449–460 (1999).
46. Fujiwara, H., Osanai, M., Matsumoto, T. & Kojima, K. K. Telomere-specific non-LTR retrotransposons and telomere maintenance in the silkworm, *Bombyx mori*. *Chromosome Res.* **13**, 455–467 (2005).
47. Biessmann, H. *et al.* Two distinct domains in *Drosophila melanogaster* telomeres. *Genetics* **171**, 1767–1777 (2005).
48. Casacuberta, E. & Pardue, M. L. HeT-A and TART, two *Drosophila* retrotransposons with a bona fide role in chromosome structure for more than 60 million years. *Cytogenet. Genome Res.* **110**, 152–159 (2005).
49. Melnikova, L. & Georgiev, P. *Drosophila* telomeres: the non-telomerase alternative. *Chromosome Res.* **13**, 431–441 (2005).
50. Biessmann, H., Kobeski, F., Walter, M. F., Kasravi, A. & Roth, C. W. DNA organization and length polymorphism at the 2L telomeric region of *Anopheles gambiae*. *Insect Mol. Biol.* **7**, 83–93 (1998).
51. Rosen, M. & Edström, J. E. DNA structures common for chironomid telomeres terminating with complex repeats. *Insect Mol. Biol.* **9**, 314–347 (2000).
52. Baudry, E. *et al.* Whole-genome scan in thelytokous-laying workers of the cape honeybee (*Apis mellifera capensis*): Central fusion, reduced recombination rates and centromere mapping using half-tetrad analysis. *Genetics* **167**, 243–252 (2004).
53. Cowan, C. R., Carlton, P. M. & Cande, W. Z. The polar arrangement of telomeres in interphase and meiosis. Rabl organization and the bouquet. *Plant Physiol.* **125**, 532–538 (2001).
54. Robertson, H. M. The *mariner* transposable element is widespread in insects. *Nature* **362**, 241–245 (1993).
55. Robertson, H. M. & MacLeod, E. G. Five major subfamilies of *mariner* transposable elements in insects, including the Mediterranean fruit fly, and related arthropods. *Insect Mol. Biol.* **2**, 125–139 (1993).
56. Robertson, H. M. & Lampe, D. J. Distribution of transposable elements in arthropods. *Annu. Rev. Entomol.* **40**, 333–357 (1995).
57. Robertson, H. M. in *Mobile DNA II* (eds Craig, N. L., Craigie, R., Gellert, M. & Lambowitz, A. M.) (ASM, Washington DC, 2002).
58. Lampe, D. J., Witherspoon, D. J., Soto-Adames, F. N. & Robertson, H. M. Recent horizontal transfer of *mellifera* subfamily mariner transposons into insect lineages representing four different orders shows that selection acts only during horizontal transfer. *Mol. Biol. Evol.* **20**, 554–562 (2003).
59. Ebert, P. R. & Hileman, J. P. t. & Nguyen, H. T. Primary sequence, copy number, and distribution of *mariner* transposons in the honey bee. *Insect Mol. Biol.* **4**, 69–78 (1995).
60. Eiglmeier, K. *et al.* Comparative analysis of BAC and whole genome shotgun sequences from an *Anopheles gambiae* region related to *Plasmodium* encapsulation. *Insect Biochem. Mol. Biol.* **35**, 799–814 (2005).
61. Goodwin, T. J., Poulter, R. T., Lorenzen, M. D. & Beeman, R. W. DIRS retroelements in arthropods: identification of the recently active TcDf1 element in the red flour beetle *Tribolium castaneum*. *Mol. Genet. Genomics* **272**, 47–56 (2004).
62. Kaminker, J. S. *et al.* The transposable elements of the *Drosophila melanogaster* euchromatin: a genomics perspective. *Genome Biol.* **3**, RESEARCH0084 (2002).
63. Gillespie, J. J., Johnston, J. S., Cannone, J. J. & Gutell, R. R. Characteristics of the nuclear (18S, 5.8S, 28S, and 5S) and mitochondrial (16S and 12S) rDNA genes of *Apis mellifera* (Insecta: Hymenoptera): Structure, organization, and retrotransposable elements. *Insect Mol. Biol.* (in the press).
64. Xiong, Y., Burke, W. D., Jakubczak, J. L. & Eickbush, T. H. Ribosomal DNA insertion elements R1Bm and R2Bm can transpose in a sequence specific manner to locations outside the 28S genes. *Nucleic Acids Res.* **16**, 10561–10573 (1988).
65. Bigot, Y., Lutchter, F., Hamelin, M. H. & Periquet, G. The 28S ribosomal RNA-encoding gene of Hymenoptera: inserted sequences in the retrotransposon-rich regions. *Gene* **121**, 347–352 (1992).
66. Krieger, M. J. & Ross, K. G. Molecular evolutionary analyses of *mariners* and other transposable elements in fire ants (Hymenoptera: Formicidae). *Insect Mol. Biol.* **12**, 155–165 (2003).
67. McAllister, B. F. & Werren, J. H. Phylogenetic analysis of a retrotransposon with implications for strong evolutionary constraints on reverse transcriptase. *Mol. Biol. Evol.* **14**, 69–80 (1997).
68. Varricchio, P. *et al.* Characterization of *Aphidius ervi* (Hymenoptera, Braconidae) ribosomal genes and identification of site-specific insertion elements belonging to the non-LTR retrotransposon family. *Insect Biochem. Mol. Biol.* **25**, 603–612 (1995).
69. Elisk, C. G. *et al.* Creating a honey bee consensus gene set. *Genome Biol.* (in the press).
70. Pearson, W. R. & Lipman, D. J. Improved tools for biological sequence comparison. *Proc. Natl Acad. Sci. USA* **85**, 2444–2448 (1988).
71. Bertone, P. *et al.* Global identification of human transcribed sequences with genome tiling arrays. *Science* **306**, 2242–2246 (2004).
72. Hillier, L. W. *et al.* Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* **432**, 695–716 (2004).
73. Zdobnov, E. M., von Mering, C., Letunic, I. & Bork, P. Consistency of genome-based methods in measuring Metazoan evolution. *FEBS Lett.* **579**, 3355–3361 (2005).
74. Bolshakov, V. N. *et al.* A comparative genomic analysis of two distant diptera, the fruit fly, *Drosophila melanogaster*, and the malaria mosquito, *Anopheles gambiae*. *Genome Res.* **12**, 57–66 (2002).
75. Zdobnov, E. M. *et al.* Comparative genome and proteome analysis of *Anopheles gambiae* and *Drosophila melanogaster*. *Science* **298**, 149–159 (2002).
76. Zdobnov, E. & Bork, P. Quantification of insect genome divergence. *Trends Genet.* (in the press).
77. Raible, F. *et al.* Vertebrate-type intron-rich genes in the marine annelid *Platynereis dumerilii*. *Science* **310**, 1325–1326 (2005).
78. Ciccarelli, F. D. *et al.* Complex genomic rearrangements lead to novel primate gene function. *Genome Res.* **15**, 343–351 (2005).
79. Tsuchimoto, M. *et al.* Conservation of novel *Mahya* genes shows the existence of neural functions common between Hymenoptera and Deuterostome. *Dev. Genes Evol.* **215**, 564–574 (2005).
80. Rubin, E. *et al.* Molecular and phylogenetic analyses reveal mammalian-like clockwork in the honey bee (*Apis mellifera*) and shed new light on the molecular evolution of the circadian clock. *Genome Res.* **16**(11), 1352–1365 (2006).
81. Velarde, R. A., Sauer, C. D., Walden, K. K., Fahrbach, S. E. & Robertson, H. M. Pteropsin: A vertebrate-like non-visual opsin expressed in the honey bee brain. *Insect Biochem. Mol. Biol.* **35**, 1367–1377 (2005).
82. Pires-daSilva, A. & Sommer, R. J. The evolution of signalling pathways in animal development. *Nature Rev. Genet.* **4**, 39–49 (2003).
83. Dearden, P. K. *et al.* Patterns of conservation and change in honeybee developmental genes. *Genome Res.* (in the press).
84. Tribolium Genome Sequencing Project. *Tribolium castaneum* v 2.0 assembly (<http://www.hgsc.bcm.tmc.edu/projects/tribolium>) (2006).
85. Lehmann, R. & Nusslein-Volhard, C. Abdominal segmentation, pole cell formation, and embryonic polarity require the localized activity of *oskar*, a maternal gene in *Drosophila*. *Cell* **47**, 141–152 (1986).
86. Schroder, R. *vasa* mRNA accumulates at the posterior pole during blastoderm formation in the flour beetle *Tribolium castaneum*. *Dev. Genes Evol.* **216**, 277–283 (2006).
87. Nakao, H. Isolation and characterization of a *Bombyx vasa*-like gene. *Dev. Genes Evol.* **209**, 312–316 (1999).

88. Dearden, P. K. Germ cell development in the Honeybee (*Apis mellifera*); *Vasa* and *Nanos* expression. *BMC Dev. Biol.* **6**, doi:10.1186/1471-213X-6-6 (17 February 2006).
89. Wilkins, A. S. *The Evolution of Developmental Pathways* (Sinauer, Sunderland, Massachusetts, 2002).
90. Davidson, E. H. *Genomic Regulatory Systems* (Academic, London, 2001).
91. Cline, T. W. & Meyer, B. J. Vive la difference: males vs females in flies vs worms. *Annu. Rev. Genet.* **30**, 637–702 (1996).
92. Albert, S., Bhattacharya, D., Klaidiny, J., Schmitzova, J. & Simuth, J. The family of major royal jelly proteins and its evolution. *J. Mol. Evol.* **49**, 290–297 (1999).
93. Drapeau, M. D., Albert, S., Kucharski, R., Prusko, C. & Maleszka, R. Evolution of the Yellow/Major Royal Jelly Protein family and the emergence of sex-specific social behavior in honeybees. *Genome Res.* (in the press).
94. Evans, J. D. & Wheeler, D. E. Differential gene expression between developing queens and workers in the honey bee, *Apis mellifera*. *Proc. Natl Acad. Sci. USA* **96**, 5575–5580 (1999).
95. Evans, J. D. & Wheeler, D. E. Expression profiles during honeybee caste determination. *Genome Biol.* **2**, RESEARCH0001 (2000).
96. Corona, M., Estrada, E. & Zurita, M. Differential expression of mitochondrial genes between queens and workers during caste determination in the honeybee *Apis mellifera*. *J. Exp. Biol.* **202**, 929–938 (1999).
97. Heppeler, C. & Hartfelder, K. Differentially expressed regulatory genes in honey bee caste development. *Naturwissenschaften* **88**, 113–116 (2001).
98. West-Eberhard, M. J. *Developmental Plasticity and Evolution* (Oxford Univ. Press, New York, 2003).
99. Colombani, J. et al. Antagonistic actions of ecdysone and insulins determine final size in *Drosophila*. *Science* **310**, 667–670 (2005).
100. Mirth, C., Truman, J. W. & Riddiford, L. M. The role of the prothoracic gland in determining critical weight for metamorphosis in *Drosophila melanogaster*. *Curr. Biol.* **15**, 1796–1807 (2005).
101. Willis, J. H., Ionomidou, V. A., Smith, R. F. & Hamodrakas, S. J. in *Comprehensive Molecular Insect Science* (eds Gilbert, L. I., Iatrou, K. & Gill, S. S.) 79–110 (Elsevier, Oxford, 2005).
102. Tellam, R. L., Wijffels, G. & Willadsen, P. Peritrophic matrix proteins. *Insect Biochem. Mol. Biol.* **29**, 87–101 (1999).
103. Jones, A. K., Raymond-Delpech, V., Thany, S. H., Gauthier, M. & Sattelle, D. B. The nicotinic acetylcholine receptor gene family of the honeybee, *Apis mellifera*. *Genome Res.* (in the press).
104. Eisenhardt, D., Kühn, C. & Lebouille, G. The PKA-CREB system encoded by the honeybee genome. *Insect Mol. Biol.* (in the press).
105. Hummon, A. B. et al. From the genome to the proteome: Uncovering peptides in the *Apis* brain. *Science* (in the press).
106. Baggerman, G., Cerstiaens, A., De Loof, A. & Schoofs, L. Peptidomics of the larval *Drosophila melanogaster* central nervous system. *J. Biol. Chem.* **277**, 40368–40374 (2002).
107. Hewes, R. S. & Taghert, P. H. Neuropeptides and neuropeptide receptors in the *Drosophila melanogaster* genome. *Genome Res.* **11**, 1126–1142 (2001).
108. Riehle, M. A., Garczynski, S. F., Crim, J. W., Hill, C. A. & Brown, M. R. Neuropeptides and peptide hormones in *Anopheles gambiae*. *Science* **298**, 172–175 (2002).
109. Hauser, F., Cazzamali, G., Williamson, M., Blenau, W. & Grimmekhuijzen, C. J. P. A review of neurohormone GPCRs present in the fruitfly *Drosophila melanogaster* and the honey bee *Apis mellifera*. *Prog. Neurobiol.* **80**, 1–19 (2006).
110. Clyne, P. J. et al. A novel family of divergent seven-transmembrane proteins: candidate odorant receptors in *Drosophila*. *Neuron* **22**, 327–338 (1999).
111. Robertson, H. M., Warr, C. G. & Carlson, J. R. Molecular evolution of the insect chemoreceptor gene superfamily in *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA* **100** (suppl. 2), 14537–14542 (2003).
112. Voshall, L. B., Amrein, H., Morozov, P. S., Rzhetsky, A. & Axel, R. A spatial map of olfactory receptor expression in the *Drosophila* antenna. *Cell* **96**, 725–736 (1999).
113. Hill, C. A. et al. G-protein-coupled receptors in *Anopheles gambiae*. *Science* **298**, 176–178 (2002).
114. Robertson, H. M. & Wanner, K. W. The chemoreceptor superfamily in the honey bee *Apis mellifera*: expansion of the odorant, but not gustatory, receptor family. *Genome Res.* (in the press).
115. Galizia, C. G. & Menzel, R. The role of glomeruli in the neural representation of odours: results from optical recording studies. *J. Insect Physiol.* **47**, 115–130 (2001).
116. Fishilevich, E. & Voshall, L. B. Genetic and functional subdivision of the *Drosophila* antennal lobe. *Curr. Biol.* **15**, 1548–1553 (2005).
117. Clyne, P. J., Warr, C. G. & Carlson, J. R. Candidate taste receptors in *Drosophila*. *Science* **287**, 1830–1834 (2000).
118. Xu, P., Atkinson, R., Jones, D. N. & Smith, D. P. *Drosophila* OBP LUSH is required for activity of pheromone-sensitive neurons. *Neuron* **45**, 193–200 (2005).
119. Foret, S. & Maleszka, R. Function and evolution of odorant binding protein gene family in a social insect, the honey bee (*Apis mellifera*). *Genome Res.* (in the press).
120. Panda, S., Hogenesch, J. B. & Kay, S. A. Circadian rhythms from flies to human. *Nature* **417**, 329–335 (2002).
121. Guillaumond, F., Dardente, H., Giguere, V. & Cermakian, N. Differential control of Bmal1 circadian transcription by REV-ERB and ROR nuclear receptors. *J. Biol. Rhythms* **20**, 391–403 (2005).
122. Morse, R. A. & Flottum, K. (eds). *Honey Bee Pests, Predators and Diseases* (A. I. Root Co., Medina, Ohio, 1997).
123. Schmid-Hempel, P. *Parasites in Social Insects* (Princeton Univ. Press, Princeton, New Jersey, 1998).
124. Evans, J. D. et al. Immune pathways and defence mechanisms in honey bees. *Insect Mol. Biol.* (in the press).
125. Aronstein, K. & Saldivar, E. Characterization of a honey bee Toll related receptor gene *Am18w* and its potential involvement in antimicrobial immune defense. *Apidologie (Celle)* **36**, 3–14 (2005).
126. Evans, J. D. Beepath: An ordered quantitative-PCR array for exploring honey bee immunity and disease. *J. Invertebr. Pathol.* **93**, 135–139 (2006).
127. Evans, J. D. & Pettis, J. S. Colony-level effects of immune responsiveness in honey bees, *Apis mellifera*. *Evol. Int. J. Org. Evol.* **59**, 2270–2274 (2005).
128. Chen, Y. P., Higgins, J. A. & Feldlaufer, M. F. Quantitative real-time reverse transcription-PCR analysis of deformed wing virus infection in the honeybee (*Apis mellifera* L.). *Appl. Environ. Microbiol.* **71**, 436–441 (2005).
129. Gregorc, A. & Bowen, I. D. Histopathological and histochemical changes in honeybee larvae (*Apis mellifera* L.) after infection with *Bacillus larvae*, the causative agent of American foulbrood disease. *Cell Biol. Int.* **22**, 137–144 (1998).
130. Bogdanov, S., Kilchenmann, V. & Imdorf, A. Acaricide residues in some bee products. *J. Apicultural. Res.* **37**, 57–67 (1998).
131. Tremolada, P., Bernardinelli, I., Colombo, M., Spreafico, M. & Vighi, M. Coumaphos distribution in the hive ecosystem: case study for modeling applications. *Ecotoxicology* **13**, 589–601 (2004).
132. Wallner, K. The use of varroacides and their influence on the quality of bee products. *Am. Bee J.* **135**, 817–821 (1995).
133. Berenbaum, M. R. in *Molecular Biology of the Toxic Response* (eds Puga, A. & Wallace, K. B.) 553–571 (Taylor & Francis, Philadelphia, 1999).
134. Oakeshott, J. G., Claudianos, C., Campbell, P. M., Newcomb, R. D. & Russell, R. J. in *Comprehensive Molecular Insect Science* (eds Gilbert, L. I., Iatrou, K. & Gill, S. S.) 309–381 (Elsevier Pergamon, Oxford, 2005).
135. Ranson, H. & Hemingway, J. in *Comprehensive Molecular Insect Science* (eds Gilbert, L. I., Iatrou, K. & Gill, S. S.) 383–402 (Elsevier Pergamon, Oxford, 2005).
136. Feyereisen, R. in *Comprehensive Molecular Insect Science* (eds Gilbert, L. I., Iatrou, K. & Gill, S. S.) 1–77 (Elsevier Pergamon, Oxford, 2005).
137. Ranson, H. et al. Evolution of supergene families associated with insecticide resistance. *Science* **298**, 179–181 (2002).
138. Claudianos, C. et al. A deficit of detoxification enzymes: Pesticide sensitivity and environmental response in the honeybee. *Insect Mol. Biol.* (in the press).
139. Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6–21 (2002).
140. Wang, Y. et al. Functional CpG methylation system in a social insect. *Science* (in the press).
141. Winston, W. M., Molodowitch, C. & Hunter, C. P. Systemic RNAi in *C. elegans* requires the putative transmembrane protein SID-1. *Science* **295**, 2456–2459 (2002).
142. Beye, M., Hartel, S., Hagen, A., Hasselmann, M. & Omholt, S. W. Specific developmental gene silencing in the honey bee using a homeobox motif. *Insect Mol. Biol.* **11**, 527–532 (2002).
143. Amdam, G. V., Simoes, Z. L., Guidugli, K. R., Norberg, K. & Omholt, S. W. Disruption of vitellogenin gene function in adult honeybees by intra-abdominal injection of double-stranded RNA. *BMC Biotechnol.* **3**, doi:10.1186/1472-6750-3-1 (20 January 2003).
144. Aronstein, K., Pankiw, T. & Saldivar, E. SID-1 is implicated in systemic gene silencing in the honey bee. *J. Apic. Res.* **45**, 20–24 (2006).
145. Thummel, C. S. From embryogenesis to metamorphosis: the regulation and function of *Drosophila* nuclear receptor superfamily members. *Cell* **83**, 871–877 (1995).
146. Velarde, R. A., Fahrback, S. & Robinson, G. E. Nuclear receptors of the honey bee: annotation and expression in the adult brain. *Insect Mol. Biol.* (in the press).
147. Gerber, S. et al. The photoreceptor cell-specific nuclear receptor gene (PNR) accounts for retinitis pigmentosa in the Crypto-Jews from Portugal (Marranos), survivors from the Spanish Inquisition. *Hum. Genet.* **107**, 276–284 (2000).
148. Kobayashi, M. et al. Identification of a photoreceptor cell-specific nuclear receptor. *Proc. Natl Acad. Sci. USA* **96**, 4814–4819 (1999).
149. Sinha, S., Ling, X., Whitfield, C. W., Zhai, C. & Robinson, G. E. Genome scan for cis-regulatory DNA motifs associated with social behavior in honey bees. *Proc. Natl Acad. Sci. USA* (in the press).
150. Whitfield, C. W. et al. Genomic dissection of behavioural maturation in the honey bee. *Proc. Natl Acad. Sci. USA* (in the press).
151. Whitfield, C. W., Cziko, A. M. & Robinson, G. E. Gene expression profiles in the brain predict behavior in individual honey bees. *Science* **302**, 296–299 (2003).
152. Savard, J. et al. Phylogenomic analysis reveals bees and wasps (Hymenoptera) at the base of the radiation of holometabolous insects. *Genome Res.* (in the press).
153. Danforth, B. N., Fang, J. & Sipes, S. Analysis of family-level relationships in bees (Hymenoptera: Apoidea) using 28S and two previously unexplored nuclear genes: CAD and RNA polymerase II. *Mol. Phylog. Evol.* **39**, 358–372 (2006).
154. Danforth, B. N., Sipes, S., Fang, J. & Brady, S. G. The history of early bee diversification based on five genes plus morphology. *Proc. Natl Acad. Sci. USA*. doi:10.1073/pnas.0604033103 (2 October 2006).
155. Marth, G. T. et al. A general approach to single-nucleotide polymorphism discovery. *Nature Genet.* **23**, 452–456 (1999).
156. Ruttner, F., Tassencourt, L. & Louveaux, J. Biometric-statistical analysis of the geographical variability of *Apis mellifera* L. I. Material and methods. *Apidologie (Celle)* **9**, 363–381 (1978).

157. Franck, P., Garnery, L., Solignac, M. & Cornuet, J. M. Molecular confirmation of a fourth lineage in honeybees from the Near East. *Apidologie (Celle)* **31**, 167–180 (2000).
158. Sheppard, W. S., Rinderer, T. E., Mazzoli, J. A., Stelzer, J. A. & Shimanuki, H. Gene flow between African- and European-derived honey bee populations in Argentina. *Nature* **349**, 782–784 (1991).
159. Iltis, H. *Life of Mendel* (W. W. Norton and Company, Inc., New York, 1923).
160. Wang, J. *et al.* SilkDB: a knowledgebase for silkworm biology and genomics. *Nucleic Acids Res.* **33**, D399–D402 (2005).
161. Robertson, H. M. Insect genomes. *Am. Entomol.* **51**, 166–171 (2005).
162. Hunt, G. *et al.* Behavioral genomics of honeybee foraging and nest defence. *Naturwissenschaften* (in the press).
163. Chimpanzee Sequencing and Analysis Consortium. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* **437**, 69–87 (2005).
164. Sarfare, S., Ahmad, S. T., Joyce, M. V., Boggess, B. & O'Tousa, J. E. The *Drosophila ninaG* oxidoreductase acts in visual pigment chromophore production. *J. Biol. Chem.* **280**, 11895–11901 (2005).
165. Ahmad, S. T., Joyce, M. V., Boggess, B. & O'Tousa, J. E. The role of *Drosophila ninaG* oxidoreductase in visual pigment chromophore biogenesis. *J. Biol. Chem.* **281**, 9205–9209 (2006).
166. Smith, W. C. & Goldsmith, T. H. Phyletic aspects of the distribution of 3-hydroxyretinal in the class insecta. *J. Mol. Evol.* **30**, 72–84 (1990).
167. Song, J., Wu, L., Chen, Z., Kohanski, R. A. & Pick, L. Axons guided by insulin receptor in *Drosophila* visual system. *Science* **300**, 502–505 (2003).
168. Shieh, B. H., Zhu, M. Y., Lee, J. K., Kelly, I. M. & Bahiraei, F. Association of INAD with NORPA is essential for controlled activation and deactivation of *Drosophila* phototransduction *in vivo*. *Proc. Natl Acad. Sci. USA* **94**, 12682–12687 (1997).
169. Townson, S. M. *et al.* Honeybee blue- and ultraviolet-sensitive opsins: cloning, heterologous expression in *Drosophila*, and physiological characterization. *J. Neurosci.* **18**, 2412–2422 (1998).
170. Skrzipek, K.-H. & Skrzipek, H. The ninth retinula cell in the ommatidium of the worker honey bee (*Apis mellifica* L.). *Z. Zellforsch. Mikrosk. Anat.* **147**, 589–593 (1974).
171. Tomlinson, A. & Struhl, G. Delta/Notch and Boss/Sevenless signals act combinatorially to specify the *Drosophila* R7 photoreceptor. *Mol. Cell* **7**, 487–495 (2001).
172. Lee, Y. *et al.* Pyrexia is a new thermal transient receptor potential channel endowing tolerance to high temperatures in *Drosophila melanogaster*. *Nature Genet.* **37**, 305–310 (2005).
173. Rosenzweig, M. *et al.* The *Drosophila* ortholog of vertebrate TRPA1 regulates thermotaxis. *Genes Dev.* **19**, 419–424 (2005).
174. Tracey, W. D. Jr, Wilson, R. I., Laurent, G. & Benzer, S. *painless*, a *Drosophila* gene essential for nociception. *Cell* **113**, 261–273 (2003).
175. O'Hagan, R., Chalfie, M. & Goodman, M. B. The MEC-4 DEG/ENaC channel of *Caenorhabditis elegans* touch receptor neurons transduces mechanical signals. *Nature Neurosci.* **8**, 43–50 (2005).
176. Si, A., Helliwell, P. & Maleszka, R. Effects of NMDA receptor antagonists on olfactory learning and memory in the honeybee (*Apis mellifera*). *Pharmacol. Biochem. Behav.* **77**, 191–197 (2004).
177. Xia, S. *et al.* NMDA receptors mediate olfactory learning and memory in *Drosophila*. *Curr. Biol.* **15**, 603–615 (2005).
178. Kucharski, R., Ball, E. E., Hayward, D. C. & Maleszka, R. Molecular cloning and expression analysis of a cDNA encoding a glutamate transporter in the honeybee brain. *Gene* **242**, 399–405 (2000).
179. Zhu, H., Yuan, Q., Froy, O., Casselman, A. & Reppert, S. M. The two CRYs of the butterfly. *Curr. Biol.* **15**, R953–R954 (2005).
180. Gaunt, M. W. & Miles, M. A. An insect molecular clock dates the origin of the insects and accords with palaeontological and biogeographic landmarks. *Mol. Biol. Evol.* **19**, 748–761 (2002).
181. Nei, M. Genetic distance between populations. *Am. Nat.* **106**, 283–292 (1972).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements Work at the BCM-HGSC was supported by grants from the NHGRI and USDA. BAC and fosmid Library construction was supported by a subcontract from a grant awarded to J.S.J., organized by D.B.W. (President, Bee Weaver Apiaries, Inc.). Fgenesh and Fgenesh++ analysis was donated by Softberry. Other support was received from NIH NIAID (H.M.R.), NRI Functional Genomics (G.E.R.) and Illinois Sociogenomics Initiative (G.E.R.), NSF (M.M.E.; B. Schatz, UIUC), Intramural Research Program NIH NLM (R.A.), USDA-NRI, California Beekeepers Assoc., Texas Beekeepers Assoc., T. W. Burleson and Son, Inc., TAMU, NIH NLM (J.G.R.), RSNZ Marsden Fund (P.K.D.), Danish Research Agency, Carlsberg Foundation, Novo Nordisk Foundation, and DFG. The authors thank the production staff at the HGSC.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to G.M.W. (gwstoc@bcm.edu).

The Honeybee Genome Sequencing Consortium

Overall project leadership: George M. Weinstock^{1,2}, Gene E. Robinson^{7,9,13,14}

Principal investigators: Richard A. Gibbs^{1,2}, George M. Weinstock^{1,2}

Community coordination: George M. Weinstock (leader)^{1,2}, Gene E. Robinson (leader)^{7,9,13,14}, Kim C. Worley (leader)^{1,2}, Jay D. Evans⁴, Ryszard Maleszka⁶, Hugh M. Robertson^{7,9,13,14}, Daniel B. Weaver¹⁶

Annotation section leaders: Martin Beye¹⁷, Peer Bork^{18,19}, Christine G. Elsik²⁰, Jay D. Evans⁴, Klaus Hartfelder²⁵, Greg J. Hunt²⁷, Hugh M. Robertson^{7,9,13,14}, Gene E. Robinson^{7,9,13,14}, Ryszard Maleszka⁶, George M. Weinstock^{1,2}, Kim C. Worley^{1,2}, Evgeny M. Zdobnov^{18,28}

Caste development and reproduction: Klaus Hartfelder (leader)²⁵, Gro V. Amdam²⁹, Márcia M. G. Bitondi²⁶, Anita M. Collins⁴, Alexandre S. Cristino³⁰, Jay D. Evans⁴, H. Michael G. Lattorff³¹, Carlos H. Lobo²⁴, Robin F. A. Moritz³¹, Francis M. F. Nunes²⁴, Robert E. Page Jr²⁹, Zilá L. P. Simões²⁶, Diana Wheeler³²

EST sequencing: Piero Carninci (leader)³³, Shiro Fukuda³³, Yoshihide Hayashizaki³³, Chikatoshi Kai³³, Jun Kawai³³, Naoko Sakazume³³, Daisuke Sasaki³³, Michihiro Tagami³³

Brain and behaviour: Ryszard Maleszka (leader)⁶, Gro V. Amdam²⁹, Stefan Albert³⁴, Geert Baggerman³⁵, Kyle T. Beggs³⁷, Guy Bloch³⁸, Giuseppe Cazzamali⁴¹, Mira Cohen³⁸, Mark David Drapeau⁴², Dorothea Eisenhardt⁴³, Christine Emore²⁷, Michael A. Ewing¹⁵, Susan E. Fahrback⁴⁸, Sylvain Forêt⁶, Cornelis J. P. Grimmelikhuijzen⁴¹, Frank Hauser⁴¹, Amanda B. Hummon¹⁵, Greg J. Hunt²⁷, Jürgen Huybrechts³⁵, Andrew K. Jones⁴⁴, Tatsuhiko Kadowaki⁵⁵, Noam Kaplan⁴⁰, Robert Kucharski⁶, Gérard Lebouille⁴³, Michal Linial^{39,40}, J. Troy Littleton⁴⁵, Alison R. Mercer³⁷, Robert E. Page Jr²⁹, Hugh M. Robertson^{7,9,13,14}, Gene E. Robinson^{7,9,13,14}, Timothy A. Richmond¹⁵, Sandra L. Rodriguez-Zas¹², Elad B. Rubin³⁸, David B. Sattelle⁴⁴, David Schlup²⁷, Liliane Schoofs³⁵, Yair Shemesh³⁸, Jonathan V. Sweedler^{13,15}, Rodrigo Velarde⁷, Peter Verleyen³⁵, Evy Vierstraete³⁵, Michael R. Williamson⁴¹

Development and metabolism: Martin Beye (leader)¹⁷, Seth A. Ament¹³, Susan J. Brown⁵⁰, Miguel Corona⁷, Peter K. Dearden³⁶, W. Augustine Dunn⁵², Michelle M. Elekonich⁵³, Christine G. Elsik²⁰, Sylvain Forêt⁶, Tomoko Fujiyuki⁵⁴, Irene Gattermeier¹⁷, Tanja Gempe¹⁷, Martin Hasselmann¹⁷, Tatsuhiko Kadowaki⁵⁵, Eriko Kage⁵⁴, Azusa Kamikouchi⁵⁴, Takeo Kubo⁵⁴, Robert Kucharski⁶, Takekazu Kunieda⁵⁴, Marcé Lorenzen⁴⁹, Ryszard Maleszka⁶, Natalia V. Milshina²⁰, Mizue Morioka⁵⁴, Kazuaki Ohashi⁵⁴, Ross Overbeek⁵⁷, Robert E. Page Jr²⁹, Hugh M. Robertson^{7,9,13,14}, Gene E. Robinson^{7,9,13,14}, Christian A. Ross⁵³, Morten Schioett¹⁷, Teresa Shippy⁵¹, Hideaki Takeuchi⁵⁴, Amy L. Toth¹⁴, Judith H. Willis⁵², Megan J. Wilson³⁶

Comparative and evolutionary analysis: Hugh M. Robertson (leader)^{7,9,13,14}, Evgeny M. Zdobnov (leader)^{18,28}, Peer Bork^{18,19}, Christine G. Elsik²⁰, Karl H. J. Gordon⁴⁶, Ivica Letunic¹⁸

Funding agency management: Kevin Hackett⁵, Jane Peterson⁵⁸, Adam Felsenfeld⁵⁸, Mark Guyer⁵⁸

Physical and genetic mapping: Michel Solignac (leader)⁵⁶, Richa Agarwala⁵⁹, Jean Marie Cornuet⁶⁰, Christine G. Elsik²⁰, Christine Emore²⁷,

Greg J. Hunt²⁷, Monique Monnerot⁵⁶, Florence Mougél⁵⁶, Justin T. Reese²⁰, David Schlipalius²⁷, Dominique Vautrin⁵⁶, Daniel B. Weaver¹⁶

Ribosomal RNA genes and related retrotransposable elements: Joseph J. Gillespie (leader)^{21,62}, Jamie J. Cannone⁶¹, Robin R. Gutell⁶¹, J. Spencer Johnston²¹

Gene prediction and consensus gene set: Christine G. Elsik (leader)²⁰, Giuseppe Cazzamali⁴¹, Michael B. Eisen^{63,64}, Cornelis J. P. Grimelikhuijzen⁴¹, Frank Hauser⁴¹, Amanda B. Hummon¹⁵, Venky N. Iyer⁶³, Vivek Iyer⁶⁵, Peter Kosarev⁶⁶, Aaron J. Mackey⁶⁷, Ryszard Maleszka⁶, Justin T. Reese²⁰, Timothy A. Richmond¹⁵, Hugh M. Robertson^{7,9,13,14}, Victor Solov'yev⁶⁸, Alexandre Souvorov⁵⁹, Jonathan V. Sweedler^{13,15}, George M. Weinstock^{1,2}, Michael R. Williamson⁴¹, Evgeny M. Zdobnov^{18,28}

Honeybee disease and immunity: Jay D. Evans (leader)⁴, Katherine A. Aronstein⁶⁹, Katarina Bilikova⁷⁰, Yan Ping Chen⁴, Andrew G. Clark⁷², Laura I. Decanini⁴, William M. Gelbart⁷³, Charles Hetru⁷⁴, Dan Hultmark⁷⁵, Jean-Luc Imler⁷⁴, Haobo Jiang⁷⁶, Michael Kanost⁵¹, Kiyoshi Kimura⁷⁷, Brian P. Lazzaro⁷¹, Dawn L. Lopez⁴, Jozef Simuth⁷⁰, Graham J. Thompson⁷⁸, Zhen Zou⁷⁶

BAC/fosmid library construction and analysis: Pieter De Jong (leader)⁷⁹, Erica Sodergren (leader)^{1,2}, Miklós Csűrös⁸⁷, Aleksandar Milosavljevic^{1,2}, J. Spencer Johnston²¹, Kazutoyo Osoegawa⁷⁹, Stephen Richards^{1,2}, Chung-Li Shu⁷⁹, George M. Weinstock^{1,2}

G+C content: Christine G. Elsik (leader)²⁰, Laurent Duret⁸⁰, Eran Elhaik²³, Dan Graur²³, Justin T. Reese²⁰, Hugh M. Robertson^{7,9,13,14}

Transposable elements: Hugh M. Robertson (leader)^{7,9,13,14}, Christine G. Elsik²⁰

Gene regulation including miRNA and RNAi: Ryszard Maleszka (leader)⁶, Daniel B. Weaver (leader)¹⁶, Gro V. Amdam²⁹, Juan M. Anzola²⁰, Kathryn S. Campbell⁷³, Kevin L. Childs²⁰, Derek Collinge⁴⁶, Madeline A. Crosby⁷³, C. Michael Dickens²⁰, Christine G. Elsik²⁰, Karl H. J. Gordon⁴⁶, L. Sian Grametes⁷³, Christina M. Grozinger⁸¹, Peter L. Jones⁹, Mireia Jorda⁸⁹, Xu Ling⁸, Beverly B. Matthews⁷³, Jonathan Miller^{1,3}, Natalia V. Milshina²⁰, Craig Mizzen¹⁷, Miguel A. Peinado⁸⁹, Justin T. Reese²⁰, Jeffrey G. Reid^{3,22}, Hugh M. Robertson^{7,9,13,14}, Gene E. Robinson^{7,9,13,14}, Susan M. Russo⁷³, Andrew J. Schroeder⁷³, Susan E. St Pierre⁷³, Ying Wang⁹, Pinglei Zhou⁷³

Superscaffold assembly: Hugh M. Robertson (leader)^{7,9,13,14}, Richa Agarwala⁵⁹, Christine G. Elsik²⁰, Natalia V. Milshina²⁰, Justin T. Reese²⁰, Daniel B. Weaver¹⁶

Data management: Kim C. Worley (leader)^{1,2}, Kevin L. Childs²⁰, C. Michael Dickens²⁰, Christine G. Elsik²⁰, William M. Gelbart⁷³, Huaiyang Jiang^{1,2}, Paul Kitts⁵⁹, Natalia V. Milshina²⁰, Justin T. Reese²⁰, Barbara Ruef⁵⁹, Susan M. Russo⁷³, Anand Venkatraman²⁰, George M. Weinstock^{1,2}, Lan Zhang^{1,2}, Pinglei Zhou⁶⁹

Chromosome structure: J. Spencer Johnston (leader)²¹, Gildardo Aquino-Perez²¹, Jean Marie Cornuet⁶⁰, Monique Monnerot⁵⁶, Michel Solignac⁵⁶, Dominique Vautrin⁵⁶

Population genetics and SNPs: Charles W. Whitfield (leader)^{7,13,14}, Susanta K. Behra⁷, Stewart H. Berlocher^{7,14}, Andrew G. Clark⁷², Richard A. Gibbs^{1,2}, J. Spencer Johnston²¹, Walter S. Sheppard⁸², Deborah R. Smith⁸³, Andrew V. Suarez^{7,11}, Neil D. Tsutsui⁸⁴, Daniel B. Weaver¹⁶, Xuehong Wei^{1,2}, David Wheeler^{1,2}

Genome assembly: George M. Weinstock (leader)^{1,2}, Kim C. Worley (leader)^{1,2}, Paul Havlak^{1,2}, Bingshan Li^{1,2}, Yue Liu^{1,2}, Erica Sodergren^{1,2}, Lan Zhang^{1,2}

(A+T)-rich DNA generation: Martin Beye (leader)¹⁷, Martin Hasselmann¹⁷, Angela Jolivet^{1,2}, Sandra Lee^{1,2}, Lynne V. Nazareth^{1,2}, Ling-Ling Pu^{1,2}, Rachel Thorn^{1,2}, George M. Weinstock^{1,2}

Tiling arrays: Viktor Stolz (leader)⁸⁵, Gene E. Robinson (leader)^{7,9,13,14}, Ryszard Maleszka⁶, Thomas Newman⁷, Manoj Samanta^{85,86}, Waraporn A. Tongprasit⁸⁵

Anti-xenobiotic defence mechanisms: Katherine A. Aronstein (leader)⁶⁹, Charles Claudianos (leader)^{6,46}, May R. Berenbaum⁷, Sunita Biswas^{6,46}, Dirk C. de Graaf⁴⁷, Rene Feyereisen⁹⁰, Reed M. Johnson⁷, John G. Oakeshott⁴⁶, Hilary Ranson⁸⁸, Mary A. Schuler¹⁰

DNA sequencing: Donna Muzny (leader)^{1,2}, Richard A. Gibbs (leader)^{1,2}, George M. Weinstock (leader)^{1,2}, Joseph Chacko^{1,2}, Clay Davis^{1,2}, Huyen Dinh^{1,2}, Rachel Gill^{1,2}, Judith Hernandez^{1,2}, Sandra Hines^{1,2}, Jennifer Hume^{1,2}, LaRonda Jackson^{1,2}, Christie Kovar^{1,2}, Lora Lewis^{1,2}, George Miner^{1,2}, Margaret Morgan^{1,2}, Lynne V. Nazareth^{1,2}, Ngoc Nguyen^{1,2}, Geoffrey Okwuonu^{1,2}, Heidi Paul^{1,2}, Stephen Richards^{1,2}, Jireh Santibanez^{1,2}, Glenford Savery^{1,2}, Erica Sodergren^{1,2}, Amanda Svatek^{1,2}, Donna Villasana^{1,2}, Rita Wright^{1,2}

Affiliations for participants: ¹Human Genome Sequencing Center, ²Department of Molecular and Human Genetics, and ³Department of Biochemistry, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030, USA. ⁴Bee Research Laboratory, BARC-E, and ⁵National Program Staff, USDA–Agricultural Research Service, Beltsville, Maryland 20705, USA. ⁶ARC Special Centre for the Molecular Genetics of Development, Visual Sciences, Research School of Biological Sciences, The Australian National University, Canberra, Australian Capital Territory 0200, Australia. ⁷Department of Entomology, ⁸Department of Computer Science, ⁹Department of Cell and Developmental Biology, ¹⁰Department of Cell and Structural Biology, ¹¹Department of Animal Biology, ¹²Animal Sciences, ¹³Neuroscience Program, ¹⁴Program in Ecology and Evolutionary Biology, and ¹⁵Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA. ¹⁶Bee Power, L.P., 16484 CR 319, Lynn Grove Road, Navasota, Texas 77868, USA. ¹⁷Heinrich-Heine Universität Düsseldorf, Institut fuer Genetik, Universitaetsstrasse 1, 40225 Duesseldorf, Germany. ¹⁸European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany. ¹⁹Max Delbrück Center for Molecular Medicine, Robert-Roessle-Strasse 10, 13125 Berlin-Buch, Germany. ²⁰Department of Animal Science, and ²¹Department of Entomology, Texas A&M University, College Station, Texas 77843, USA. ²²Department of Chemistry, and ²³Department of Biology and Biochemistry, University of Houston, Houston, Texas 77204, USA. ²⁴Departamento de Genética, Faculdade de Medicina de Ribeirão Preto, ²⁵Departamento de Biologia Celular e Molecular e Bioagentes Patogênicos, Faculdade de Medicina de Ribeirão Preto, and ²⁶Departamento de Biologia, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto 14049-900, Brazil. ²⁷Department of Entomology, Purdue University, West Lafayette, Indiana 47907, USA. ²⁸Department of Genetic Medicine and Development, University of Geneva Medical School CMU, 1 rue Michel-Servet, 1211 Geneva, Switzerland. ²⁹School of Life Sciences, Arizona State University, PO Box 874501, Tempe, Arizona 85287-4501, USA. ³⁰Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, Brazil. ³¹Institut für Zoologie, Molekulare Ökologie, Martin-Luther-Universität Halle-Wittenberg, Hoher Weg 4, D-06099 Halle (Saale), Germany. ³²Department of Entomology, University of Arizona, Tucson, Arizona 85721-0036, USA. ³³Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center, Yokohama 230-0045, Japan. ³⁴Institut für Medizinische Strahlenkunde und Zellforschung, Versbacher Strasse 5, 97078 Würzburg, Germany. ³⁵Laboratory of Developmental Physiology, Genomics and Proteomics, K.U. Leuven, Naamsestraat 59 B-3000 Leuven, Belgium. ³⁶Laboratory for Evolution and Development, Biochemistry Department, and ³⁷Zoology Department, University of Otago, PO Box 56, Dunedin, New Zealand. ³⁸Department of Evolution, Systematics, and Ecology, ³⁹The Sudarsky Center for Computational Biology, and ⁴⁰Department of Biological Chemistry, The Alexander Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem 91904, Israel. ⁴¹Center for Functional and Comparative Insect Genomics, Department of Cell Biology and Comparative Zoology, Institute of Biology, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen, Denmark. ⁴²Department of Biology, New York University, New York, New York 10003, USA. ⁴³Neurobiology, FB Biology/Chemistry/Pharmacy, Free University Berlin, Koenigin-Luise-Strasse 28/30, 14195 Berlin, Germany. ⁴⁴MRC Functional Genetics Unit, Department of Physiology Anatomy and Genetics, Le Gros Clark Building, University of Oxford, South Parks Road, Oxford OX1 3QX, UK. ⁴⁵The Picower Institute for Learning and Memory and the Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. ⁴⁶CSIRO Entomology, GPO Box 1700, Canberra, Australian Capital Territory 2601, Australia. ⁴⁷Laboratory of Zoophysiology, University of Ghent, K. L. Ledeganckstraat 35, B-9000 Ghent, Belgium. ⁴⁸Department of Biology, Wake Forest University, Winston-Salem, North Carolina 27109, USA. ⁴⁹USDA-ARS-GMPRC, 1515 College Avenue, Manhattan, Kansas 66502, USA. ⁵⁰Division of Biology, Ackert Hall, ⁵¹Department of Biochemistry, Kansas State University, Manhattan, Kansas 66506, USA. ⁵²Department of Cellular Biology, University of Georgia, Athens, Georgia 30602, USA. ⁵³School of Life Sciences, University of Nevada Las Vegas, 4505 Maryland Parkway, Box 454004, Las Vegas, Nevada 89154-4004, USA. ⁵⁴Graduate School of Science, The University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan. ⁵⁵Graduate School of Biocultural Sciences, Nagoya University, Chikusa, Nagoya 464-8601, Japan. ⁵⁶Laboratoire Evolution, Génomes et Spéciation Centre National de la Recherche Scientifique, 91198 Gif-sur-Yvette, France. ⁵⁷Fellowship for Interpretation of Genomes, 15W155 81st Street, Burr Ridge, Illinois 60527, USA. ⁵⁸US National Institutes of Health, National Human Genome Research Institute, 31 Center Drive, Bethesda, Maryland 20892, USA. ⁵⁹National Center for Biotechnology Information, National Library of Medicine, Department of Health and Human Services, 8600 Rockville Pike, Bethesda, Maryland 20894, USA. ⁶⁰Centre de Biologie et de Gestion des Populations, Institut National de la Recherche

Agronomie, 34988 Saint-Gély-du-Fesc, France.⁶¹Institute for Cellular and Molecular Biology and Section of Integrative Biology, University of Texas, Austin, Texas 78712, USA.
⁶²Virginia Bioinformatics Institute 0477, Bioinformatics Facility, Washington Street, Virginia Tech, Blacksburg, Virginia 24061, USA.⁶³Division of Genetics, Genomics and Development, Department of Molecular and Cell Biology, University of California, Berkeley, California 94720, USA.⁶⁴Genomics Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA.⁶⁵The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK.⁶⁶Softberry Inc., 116 Radio Circle, Suite 400, Mount Kisco, New York 10549, USA.⁶⁷Penn Genomics Institute, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA.⁶⁸Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK.⁶⁹Honey Bee Unit, USDA-ARS, 2413 East highway 83, Number 213, Weslaco, Texas 78596, USA.⁷⁰Slovak Academy of Sciences, Dubravska cesta 21, 845 51 Bratislava 45, Slovakia.⁷¹Department of Entomology, and ⁷²Department of Molecular Biology and Genetics, Cornell University, Ithaca 14853, New York, USA.⁷³Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts 02138, USA.⁷⁴Institut de Biologie Moléculaire et Cellulaire, CNRS, 15 rue René Descartes, 67084 Strasbourg Cedex, France.⁷⁵Umeå Centre for Molecular Pathogenesis, By. 6L, Umeå University, S-901 87 Umeå, Sweden.⁷⁶Department of Entomology and Plant Pathology, Oklahoma State University, 127 NRC, Stillwater, Oklahoma 74078, USA.⁷⁷National Institute of Livestock and Grassland Science, 3-1-1 Kannondai, Tsukuba, Ibaraki, 305-8517, Japan.⁷⁸School of Biological Sciences, University of Sydney, New South Wales 2006, Australia.⁷⁹BACPAC Resources, Children's Hospital Oakland Research Institute, Oakland, California 94609, USA.
⁸⁰Laboratoire de Biométrie et Biologie Evolutive, UMR 5558, CNRS, Univ. Lyon 1, 69622 Villeurbanne Cedex, France.⁸¹Department of Entomology, W.M. Keck Center for Behavioral Biology, Gardner Hall, MC 7613, North Carolina State University, Raleigh, North Carolina 27695, USA.⁸²Department of Entomology, Washington State University, Pullman, Washington 99164, USA.⁸³Department of Ecology & Evolutionary Biology/Entomology, Haworth Hall, 1200 Sunnyside Avenue, University of Kansas, Lawrence, Kansas 66045, USA.⁸⁴Department of Ecology and Evolutionary Biology, University of California, Irvine, 321 Steinhaus Hall, Irvine, California 92697, USA.⁸⁵NASA Ames Genome Research Facility, Moffet Field, California 94035, USA.⁸⁶Systemix Institute, Cupertino, California 95014, USA.⁸⁷Département d'informatique et de recherche opérationnelle, Université de Montréal, CP 6128 succ. Centre-Ville, Montréal, Québec H3C 3J7, Canada.⁸⁸Vector Research, Liverpool School of Tropical Medicine, Pembroke Place, Liverpool L3 5QA, UK.⁸⁹Research Institute of Oncology, L'Hospitalet 08907, Catalonia, Spain.⁹⁰Institut National de la Recherche Agronomique and Université de Nice Sophia Antipolis, UMR 1112, Centre de Recherche de Sophia Antipolis, 06903 Sophia Antipolis, France.

ARTICLES

Symbiosis insights through metagenomic analysis of a microbial consortium

Tanja Woyke^{1,2}, Hanno Teeling³, Natalia N. Ivanova¹, Marcel Huntemann³, Michael Richter³, Frank Oliver Gloeckner^{3,4}, Dario Boffelli^{1,2}, Iain J. Anderson¹, Kerrie W. Barry¹, Harris J. Shapiro¹, Ernest Szeto¹, Nikos C. Kyrpides¹, Marc Musmann³, Rudolf Amann³, Claudia Bergin³, Caroline Ruehland³, Edward M. Rubin^{1,2} & Nicole Dubilier³

Symbioses between bacteria and eukaryotes are ubiquitous, yet our understanding of the interactions driving these associations is hampered by our inability to cultivate most host-associated microbes. Here we use a metagenomic approach to describe four co-occurring symbionts from the marine oligochaete *Olavius algarvensis*, a worm lacking a mouth, gut and nephridia. Shotgun sequencing and metabolic pathway reconstruction revealed that the symbionts are sulphur-oxidizing and sulphate-reducing bacteria, all of which are capable of carbon fixation, thus providing the host with multiple sources of nutrition. Molecular evidence for the uptake and recycling of worm waste products by the symbionts suggests how the worm could eliminate its excretory system, an adaptation unique among annelid worms. We propose a model that describes how the versatile metabolism within this symbiotic consortium provides the host with an optimal energy supply as it shuttles between the upper oxic and lower anoxic coastal sediments that it inhabits.

Symbiosis has a major role in shaping the evolution and diversity of eukaryotic organisms¹. Remarkably, it is only in recent times that there has been an emerging recognition that most eukaryotic organisms are intimately associated with a complex community of beneficial microbes that are essential for their development, health and interactions with the environment². This renaissance in symbiosis research stems from advances in molecular approaches that have enabled the study of natural microbial consortia using cultivation-independent methods^{3–5}. Metagenomic analyses have provided a new dimension in the study of community organization and metabolism in natural microbial communities^{6–10}. So far, however, genomic analyses of symbiotic microbes from eukaryotes have been confined to individual species (for the only exception, see ref. 11), limiting our ability to understand the intricate interactions involving communication, competition and resource partitioning that shape symbiotic microbial communities.

Here we use random shotgun sequencing and nucleotide-signature-based binning to study the symbiotic community in *Olavius algarvensis*. This marine worm belongs to a group of oligochaetes (phylum Annelida) that lack a mouth, gut and anus, and are unique among annelid worms in having reduced their nephridial excretory system¹². They live in obligate and species-specific associations with multiple extracellular bacterial endosymbionts located just below the worm cuticle¹². As the symbionts have yet to be grown in culture, their phylogeny has only been accessible through 16S ribosomal RNA analysis and fluorescence *in situ* hybridization (FISH)^{13,14}. *O. algarvensis* lives in coastal Mediterranean sediments and harbours a chemoautotrophic sulphur-oxidizing gammaproteobacterium ($\gamma 1$ symbiont) and a deltaproteobacterial sulphate reducer ($\delta 1$ symbiont), recently shown to be engaged in an endosymbiotic sulphur cycle¹⁴. An additional gamma- and delta-proteobacterial symbiont ($\gamma 3$ and $\delta 4$ symbionts) of unknown function occur consistently in these hosts, and, in some individuals, a spirochaete has been observed

as a minor part of the symbiotic consortium (see Supplementary Fig. 1a)¹².

Given that most chemosynthetic hosts harbour only one or two bacteria, the association of gutless oligochaetes with multiple symbionts is remarkable, raising a series of questions about their interactions with each other, their host and the environment. How does the symbiosis compensate for the loss of digestive and excretory systems in the host, what do the various partners gain from this relationship, and is it mutually obligate? What is the selective advantage for *O. algarvensis* in harbouring multiple symbiotic partners? Our metagenomic analyses explain how the bacterial consortium meets the energy and waste management needs of its oligochaete host. We describe how resource partitioning between the phylogenetically diverse symbionts benefits both the symbionts and the worm in the heterogeneous environment. Finally, we propose a model showing that the selective advantage of harbouring multiple symbionts lies in their ability to supply their host with energy from an abundant and diverse supply of reducing equivalents and electron acceptors as it shuttles between the oxidized and reduced sediment layers.

Metagenomic data analysis and binning

Pooled samples of 200 *O. algarvensis* specimens per library were shotgun-sequenced (see Supplementary Information, Supplementary Fig. 1b) and the sequences assembled using the whole-genome shotgun assembler JAZZ¹⁵ (Supplementary Methods, Supplementary Fig. 2). To assign the metagenomic scaffolds to their phylotype origin, we used a combinatory binning approach based on intrinsic DNA signatures (see Supplementary Methods). Binning of the *Olavius* symbionts' metagenome resulted in the formation of four distinct clusters (Fig. 1). The presence of the corresponding rRNA operons, which represented the only rRNA operons found in these bins and the assembly, enabled us to identify them as the *O. algarvensis* symbionts $\delta 1$, $\delta 4$, $\gamma 1$ and $\gamma 3$ (Fig. 1, Supplementary Table 1;

¹DOE Joint Genome Institute, Walnut Creek, California 94598, USA. ²Lawrence Berkeley National Laboratory, Genomics Division, Berkeley, California 94720, USA. ³Max Planck Institute for Marine Microbiology, 28359 Bremen, Germany. ⁴International University Bremen, 28759 Bremen, Germany.

Supplementary Table 2 provides a comparison to other symbiont genomes). This study illustrates the usefulness of our nucleotide frequency method for the assignment of metagenomic scaffolds from a natural microbial community to a phylotype when using shotgun sequencing. It would not have been possible to reconstruct genome assemblies of the four symbionts purely on the basis of GC contents and scaffold read depth, and without a closely related, fully sequenced reference genome.

Symbiont bin assignments based on 16S rRNA genes were confirmed by phylogenetic analysis of predicted proteins within each cluster of scaffolds (Supplementary Fig. 3), and the clusters were furthermore supported by the distribution of 49 single-copy genes. Although the populations are not clonal, the frequencies of polymorphic sites found in the four symbiont bins, ranging from 0.01–0.1% (Supplementary Table 4), were rather low compared with environmental microbes such as those from acid mine drainage⁸. The following discussion describing the metabolism of the symbionts is based on the genes found in each symbiont bin. With a focus on capturing the core metabolic pathways present in the community, it is of minor importance if the genes within each bin originated from a single strain or represent a pan-genome of several very closely related strains¹⁶.

Carbon and energy metabolism

Gammaproteobacterial symbionts. Chemoautotrophic symbionts feed their hosts by providing them with organic carbon from

autotrophic CO₂ fixation driven by oxidation of reduced inorganic compounds such as sulphide. In agreement with previous studies indicating the chemoautotrophic, sulphur-oxidizing nature of the *O. algarvensis* γ 1 symbiont¹², our analysis of the γ 1 bin revealed the presence of genes required for autotrophic CO₂ fixation via the Calvin–Benson–Bassham cycle using type I ribulose 1,5-bisphosphate carboxylase–oxygenase (Rubisco; *cbb*), the oxidation of reduced sulphur compounds (genes such as *dsr*, *fcc* and *sox*), and the storage of sulphur in globules (*sgpB*, encoding one of three known sulphur-globule proteins).

Unexpectedly, our metagenomic analyses revealed that the γ 3 symbiont of *O. algarvensis* is also a sulphur-oxidizing chemoautotroph. Several gutless oligochaete species are known to harbour γ 3 symbionts but the metabolism of these bacteria was previously unknown and the benefit of harbouring additional Gammaproteobacteria is unclear. The nearly complete genomic sequence for the *O. algarvensis* γ 3 symbiont obtained in this study is notable, as this is the only genome from a chemoautotrophic symbiont sequenced thus far. The γ 3 bin carries all the genes required for a thiotrophic (sulphur-oxidizing) metabolism including those needed for the oxidation of reduced sulphur compounds (including *dsr*, *apr*, *sat*, *fcc* and *sox*) as well as autotrophic CO₂ fixation by means of genes closely related to but phylogenetically distinct from the γ 1 symbiont (Fig. 2). The absence of genes encoding sulphur-globule proteins in the near-complete γ 3 bin indicates that these symbionts do not store sulphur, supporting transmission electron microscopy analyses showing that only γ 1 symbionts contain sulphur globules (*O. Giere*, personal communication). In addition to using oxygen as an electron acceptor, the presence of *nap* and *nir* gene clusters indicates that the γ 3 symbionts couple oxidation of reduced sulphur compounds to dissimilatory nitrate reduction under oxygen-limiting conditions (Fig. 2). In deeper sediment layers, with neither oxygen nor nitrate, both Gammaproteobacteria have the ability to use fumarate as an electron acceptor for the oxidation of reduced sulphur compounds (Fig. 2).

Deltaproteobacterial symbionts. The presence of genes characteristic of dissimilatory sulphate reduction (such as *dsr*, *qmo* and *apr*) in both the δ 1 and δ 4 bins indicates that these symbionts are sulphate-reducing bacteria that use oxidized sulphur compounds such as sulphate as an electron acceptor, thereby producing sulphide (Fig. 2). In addition to the syntrophic cycling of sulphate and sulphide between the gamma- and delta-proteobacterial *O. algarvensis* symbionts¹⁴, intermediate sulphur compounds such as tetrathionate and thiosulphate may also be cycled between the symbionts. The δ 4 symbiont seems to be able to reduce sulphur compounds of intermediate oxidation states, as suggested by the presence of a multi-haem cytochrome most closely related to tetrathionate reductase of *Shewanella oneidensis*¹⁷ and located in a chromosomal cluster with molybdopterin-dependent dehydrogenase related to thiosulphate reductase of *Wolinella succinogenes*. Cycling of intermediate sulphur compounds is energetically more favourable than the exchange of sulphide and sulphate, as shown previously in experiments with free-living sulphate reducers and sulphur oxidizers¹⁸.

Heterotrophy is important in sulphate-reducing bacteria and correspondingly we found in the δ 1 bin genes coding for the transport and utilization of a large variety of carbohydrate substrates, including uronic acids (glucuronate, galacturonate and fructuronate), xylose, fructose, dihydroxyacetone and polyols (mannitol, sorbitol and glycerol). Although all sulphate-reducing bacteria are heterotrophic only some fix CO₂, and it is intriguing that both deltaproteobacterial symbionts are capable of autotrophic carbon fixation via the reductive acetyl-coenzyme-A (CoA) pathway, as well as via the reductive tricarboxylic acid (TCA) cycle (Fig. 2). Thus, *O. algarvensis* has established an association with four symbionts that are all capable of providing it with organic carbon through three different autotrophic pathways.

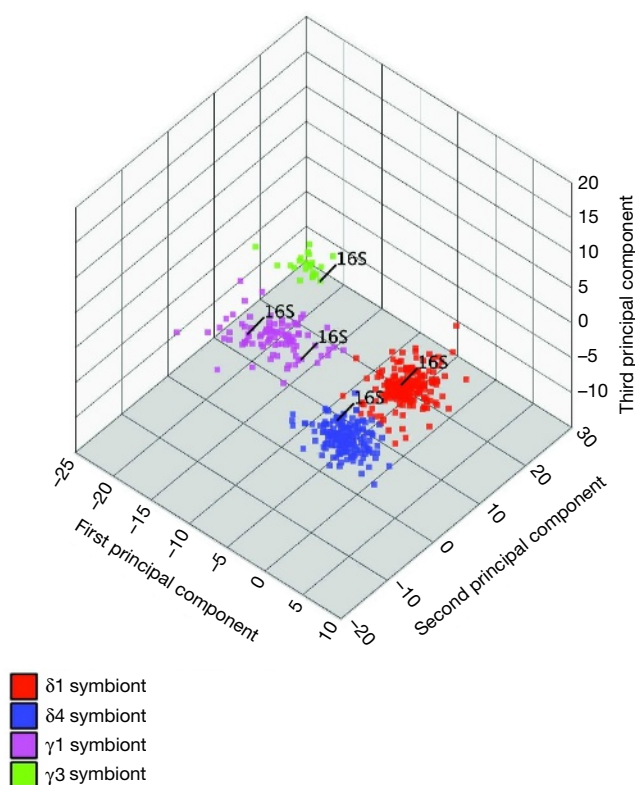


Figure 1 | Clustering of the *O. algarvensis* symbiont scaffolds. Visualization of the first three components of a principal component analysis, in which GC-content, net-read-depth z-scores for all possible 64 trinucleotides and 256 tetranucleotides were incorporated with equal weight (z-scores calculated with TETRA³⁹, and normalized by length). The colours represent the four clusters of scaffolds (calculated with MetaClust³⁵) that were binned based on GC-content, dinucleotide relative abundance, Markov-model-based statistical evaluations of tri-, tetra- and penta-mer over- and under-representation, and normalized chaos game representations for tri- to hexamers; sequences <5 kilobases (kb) in length are not represented. Scaffolds containing 16S rRNA genes are indicated.

One of the most common electron donors for autotrophic sulphate-reducing bacteria is hydrogen. We found gene clusters encoding periplasmic Ni-Fe hydrogenases, a transmembrane high-molecular-weight cytochrome *c* (hmc) complex, and a type II tetrahaem cytochrome *c*₃ (TpII-*c*₃) in the bins of both sulphate-reducing symbionts, as well as a type I tetrahaem cytochrome *c*₃ (TpI-*c*₃) in the $\delta 1$ bin (Fig. 2). This is a compelling indication for the uptake and oxidation of molecular hydrogen using sulphate as an electron acceptor¹⁹. It is not clear if hydrogen is provided by the γ -symbionts. Within the $\gamma 3$ bin, we found genes encoding a pyruvate ferredoxin oxidoreductase (POR), typically used in an alternative route for

pyruvate oxidation²⁰ and indicative of hydrogen release from low-potential ferredoxins. Released hydrogen could subsequently be taken up by the sulphate-reducing symbionts leading to hydrogen syntrophy within the microbial consortium. Alternative electron donors to hydrogen include glycerol, lactate, proline and betaine, and potentially glycolate and other 2-hydroxy acids, as well as succinate, acetate and propionate.

Symbiont-host interactions

'Feeding' of the host. In other chemosynthetic associations, the symbionts provide their hosts with nutrition using either reduced

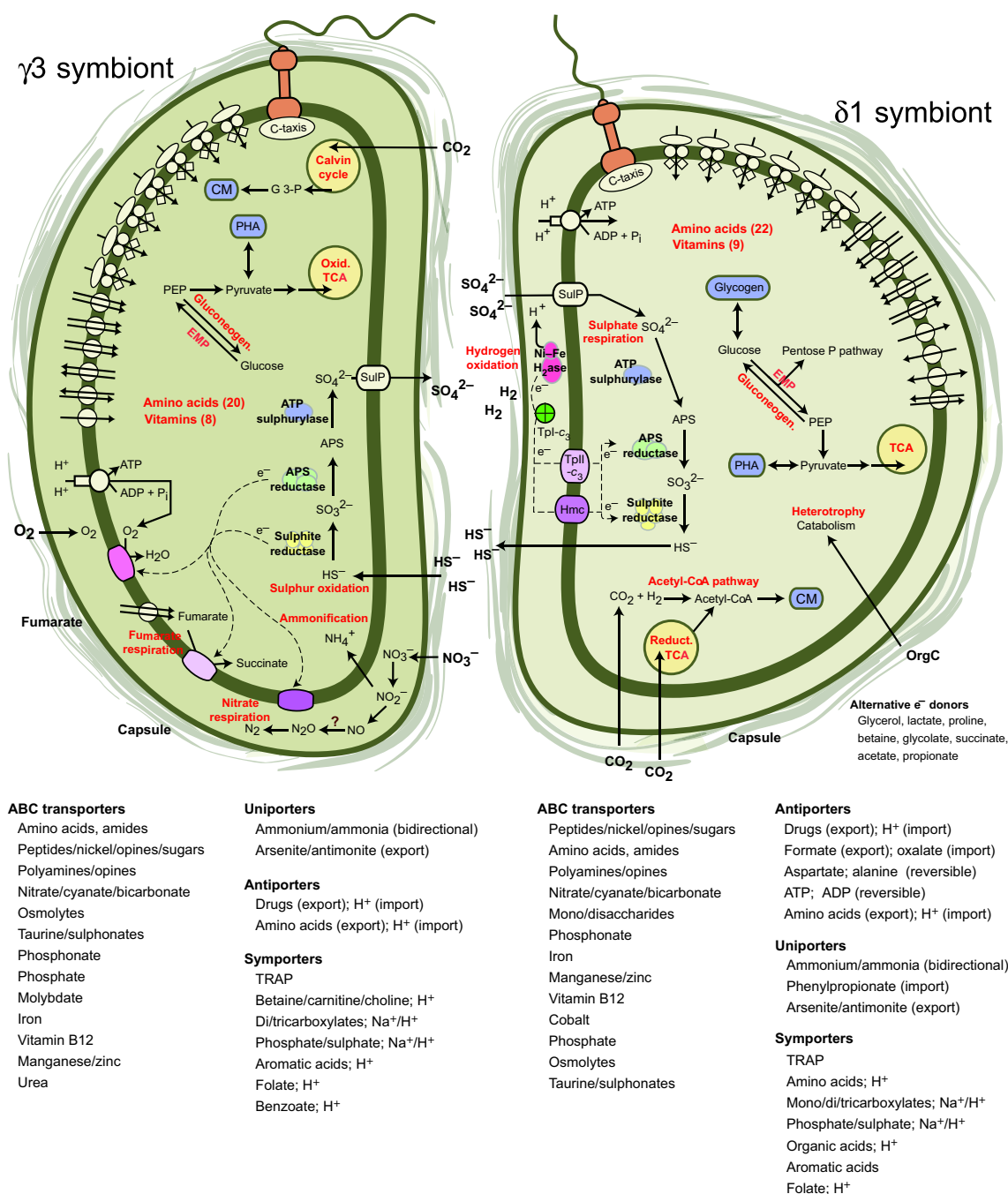


Figure 2 | Reconstruction of the symbionts' physiology. APS, adenosine 5'-phosphosulphate; CM, cell material; CoA, coenzyme A; C-taxis, chemotaxis; EMP, Embden-Meyerhof pathway; G 3-P, glyceraldehyde 3-phosphate; H₂ase, hydrogenase; Hmc, high-molecular-weight cytochrome *c*; OrgC, organic compounds; PEP, phosphoenolpyruvate; PHA, polyhydroxyalkanoates; TCA, tricyclic acid; TpI/II-*c*₃, type I/II tetrahaem

cytochrome *c*₃; TRAP, tripartite ATP-independent periplasmic. SulP, sulphate permease. A question mark (?) indicates the lack of nitric oxide reductase in the $\gamma 3$ genome bin; numbers in parentheses indicate the numbers of amino acids/vitamin biosynthesis pathways found (Supplementary Table 5).

sulphur compounds or methane as their energy source²¹. In the *O. algarvensis* symbiosis, both reduced sulphur compounds and hydrogen can be used as energy sources, and all four symbionts have the potential to fix CO₂ into organic carbon via autotrophy. In addition, the sulphate-reducing symbionts can feed the oligochaete host through heterotrophy by taking up dissolved organic carbon compounds from the environment. Almost all amino acids and a variety of vitamins can be synthesized by the symbionts to provide their host with these required nutrients (Supplementary Table 5). Nutrient transfer to the host is likely to occur via intracellular uptake and digestion of the bacteria, as the number of genes encoding amino-acid exporters was not elevated in the symbionts when compared to those of free-living bacteria, and the only known family of sugar exporters²² was not encoded in any of the symbiont bins. This conclusion is supported by morphological analyses showing symbiont lysis in the basal region of the worm's epidermis²³.

Host waste recycling. The reduction of nephridia in the oligochaete host, used for the excretion of nitrogenous waste compounds and osmoregulation, suggests that its symbionts have adapted to carry out these functions. Most aquatic organisms excrete ammonium, but urea—which also functions as a common organic osmolyte—has also been found in marine worms with symbionts, such as the hydrothermal vent worm *Riftia pachyptila*²⁴. The genome bins for the $\delta 1$, $\delta 4$ and $\gamma 3$ symbionts encode a bidirectional uniporter (Amt family transporter²⁵) for the counter-ion-independent and energy-independent uptake of ammonium. A probable urea ABC transporter for urea uptake is present in the $\gamma 3$ bin, adjacent to a urease operon that encodes genes involved in urea hydrolysis. Ammonium and urea uptake by the symbionts would not only aid the host in the removal of these toxic waste products, but also lead to the conservation of valuable nitrogen by the symbionts (Fig. 3).

Marine invertebrates are typical osmoconformers, maintaining their cell volume primarily with organic osmoregulatory compounds such as amino acids, taurine, glycine betaine, trimethylamine *N*-oxide (TMAO) and urea, as well as polyols and sugars²⁶. The

presence of gene clusters encoding proteins used for taurine and glycine betaine import and catabolism in the $\gamma 3$ bin may indicate the use of these osmolytes as both a carbon and nitrogen source²⁷ (Fig. 3). Furthermore, we found pathways for TMAO degradation in the *O. algarvensis* symbionts. Microbial TMAO degradation includes its conversion to trimethylamine by TMAO reductase, and further demethylation of trimethylamine either by trimethylamine and dimethylamine dehydrogenases found in Bacteria²⁸ or by trimethylamine, dimethylamine and monomethylamine methyltransferases found mostly in Archaea²⁹. Genes coding for the key enzymes in both pathways were present in the $\gamma 3$ and $\delta 1$ symbionts, suggestive of their TMAO catabolic activity (Fig. 3). The availability of the osmolyte TMAO would, furthermore, be particularly advantageous for the sulphur-oxidizing symbionts that could use this organic carbon compound as an alternate electron acceptor in the absence of oxygen or nitrate. Although the $\gamma 3$ bin encodes enzymes likely to be involved in the reduction of TMAO, their specificity for this substrate could not be clearly identified.

Polyamines are essential in all organisms for DNA stabilization, DNA replication and cell proliferation³⁰, and represent additional products of host protein breakdown. We found abundant gene clusters encoding ABC transporters for the uptake of the polyamines putrescine and spermidine in the $\delta 1$ bin, and putrescine in the $\gamma 3$ bin (Fig. 3).

Finally, evidence is provided for the recycling of host fermentation waste products such as the dicarboxylate succinate, as well as the monocarboxylates acetate and propionate. Pathways for their utilization and a variety of potential dicarboxylate transporters, and probable monocarboxylate transporters, were found in all four symbiont bins. The $\delta 1$ bin encodes 23 tripartite ATP-independent periplasmic (TRAP)-T family dicarboxylate transporters³¹, some of which are likely to be involved in monocarboxylate and dicarboxylate transport (Fig. 3).

A mutually obligate relationship?

The lack of a digestive and excretory system in *O. algarvensis* means that its symbionts are crucial for its survival. But is this relationship mutually obligate or can the symbionts survive outside of the host in a free-living stage? Several pieces of evidence from the sequence data and the metabolic reconstruction analyses suggest a lack of evidence for obligate host-dependence of any of the symbionts. This includes the observation that the genomes of the extracellular bacteria do not show AT-bias or genome size reduction (Supplementary Table 2) and there was no obvious loss of essential metabolic pathways in the $\delta 1$ or $\gamma 3$ bins suggestive of host-dependence, as is the case for many obligate host-associated bacteria^{32,33}. Finally, we found genes required for cell motility via a flagellum in the $\delta 1$, $\delta 4$ and $\gamma 3$ bins. Although this evidence suggests that the symbionts associated with the worm may have a free-living stage, the presence of a remarkable number of transposases in the $\gamma 3$ and $\gamma 1$ symbiont bins (7.5% and 20.5%, respectively; Supplementary Tables 4 and 6) indicates that these symbionts may be in transition to an obligate symbiotic lifestyle. Bacteria that have recently evolved into obligate symbionts show an increase in frequency of mobile elements, representing a source for chromosomal rearrangements and gene inactivation³⁴ (for symbiont transmission hypotheses, see Supplementary Information).

From the metagenome to the environment

The oligochaete symbiosis is inseparably linked to the geochemical properties of the environment, needing access to both reduced and oxidized compounds for energy production and net carbon fixation (Fig. 4). In the upper oxic sediment layers of the *O. algarvensis* habitat, where no sulphide is present, both the $\gamma 1$ and the $\gamma 3$ symbiont can use reduced sulphur compounds produced internally by the sulphate-reducing symbionts in a syntrophic sulphur cycle. (We have shown previously that the sulphate-reducing symbionts can produce sulphide in *O. algarvensis* under microaerobic conditions comparable

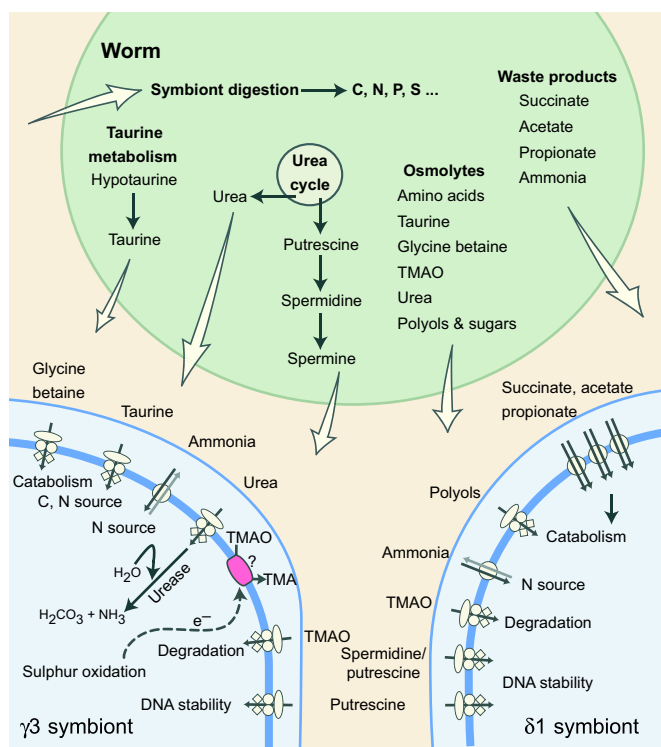


Figure 3 | Reconstruction of symbiont–host interactions. The metagenomic data uncovered pathways for the uptake and recycling of organic osmolytes and excretion products of the worm by its symbionts. TMA, trimethylamine; TMAO, trimethylamine *N*-oxide.

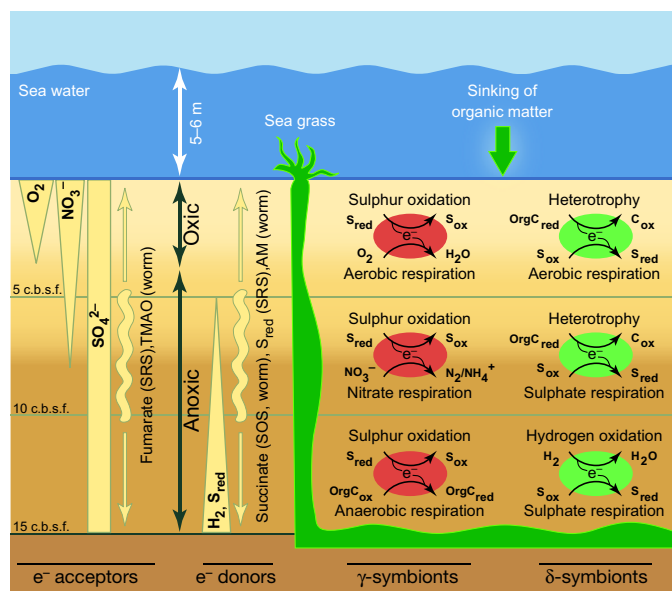


Figure 4 | Model for energy metabolism in the symbiosis. *O. algarvensis* inhabits shallow Mediterranean ocean sediments (5–15-cm depth). Electron acceptors and donors are available to the symbionts in all sediment layers, with some supplied from the environment (shown in the yellow rectangle/triangles) and some internally (shown next to the yellow worms). Carbon is gained autotrophically using reduced sulphur compounds (γ -symbionts) or hydrogen (δ -symbionts), as well as heterotrophically (δ -symbionts). AM, anaerobic metabolites; c.b.s.f., centimetres below sea floor; OrgC, organic compounds; S_{ox} , oxidized sulphur compounds; S_{red} , reduced sulphur compounds; SOS, sulphur-oxidizing symbionts; SRS, sulphate-reducing symbionts; TMAO, trimethylamine *N*-oxide.

to those in the lower oxic zone at 2–5-cm sediment depth¹⁴.) The γ 1 symbiont can also gain energy independently of the sulphate-reducing symbionts by oxidizing the large supply of sulphur stored in its cytoplasm. For both gammaproteobacterial symbionts, oxygen would be the energetically most-favourable electron acceptor (Fig. 4).

As the worm migrates downwards, it encounters sediment layers in which oxygen is no longer present. Under these conditions, the γ 3 symbiont can use nitrate from the environment for the oxidation of reduced sulphur compounds. Given the extremely low concentrations of sulphide in this layer as well as in the deeper reduced sediment layers of the Elba habitat (in the low nM range¹⁴), it is likely that the sulphate-reducing symbionts provide a large part of the reduced sulphur compounds for the γ -symbionts under most conditions.

In the deeper sediment layers characterized by reducing conditions and the absence of oxygen or nitrate, hydrogen oxidation by the sulphate-reducing symbionts may occur, in which hydrogen is used as an energy source for the autotrophic fixation of inorganic carbon. As hydrogen concentrations are commonly very low in most marine sediments, heterotrophic pathways should also have an important role for the deltaproteobacterial symbionts under these reducing conditions. Although energetically less favourable than nitrate or oxygen, organic electron acceptors including TMAO (host-derived) and fumarate (produced by the δ -symbionts) are provided internally within the symbiosis and may be used by the γ -symbionts for the oxidation of reduced sulphur compounds. This would enable the γ 1 symbiont to replenish its sulphur stores, which could be fully oxidized using the energetically more favourable electron acceptor oxygen when the worm returns to the oxic zone. Fumarate respiration by the γ -symbionts would produce succinate, which could be used by the deltaproteobacterial symbionts as an energy source and reoxidized to fumarate, thus leading to a syntrophic cycling of reductants and oxidants.

Our analysis of the *O. algarvensis* microbial symbiont genomes has provided insights into how resources are used and shared among the

different symbionts and with their host, and how different metabolic pathways are used by the symbionts to generate energy as the worm migrates through the chemocline. We have shown how the *O. algarvensis* symbiosis is unique among known chemosynthetic associations, as reductants and oxidants are not only supplied from the environment but also internally to drive energy production. Thus, this comprehensive metagenomic analysis shows that these highly integrated synergistic assemblages of multiple bacterial partners provide their eukaryotic host with an optimal energy supply and waste management through resource partitioning and cooperation during syntrophic cycling of oxidized and reduced compounds.

METHODS

Further details for all methods used in this study are provided in Supplementary Information.

O. algarvensis specimens were collected off Capo di Sant' Andrea, Elba, Italy. Metagenomic libraries were constructed from 200 pooled *O. algarvensis* specimens per library. A small-insert pMCL200 library was made from a nycodenz-separated, symbiont-enriched sample and two pCC1Fos fosmid libraries were constructed using the CopyControl Fosmid Library Construction kit (Epicentre). 16S rRNA polymerase chain reaction (PCR) libraries were created from the DNA sources used for each of the libraries and approximately 384 clones sequenced and analysed. From the shotgun libraries, we created 204 megabases (Mb) of vector- and quality-trimmed sequence. This data was assembled using JAZZ, resulting in a set of 2,286 scaffolds, which were binned using a combinatorial approach based on dimer-to-hexamer frequencies using the newly developed MetaClust program³⁵. Final clusters with 511 scaffolds were verified by phylogenetic affiliation of each scaffold based on the most common phylogeny of its predicted proteins using a bayesian classifier, and by checking for paralogues of 49 genes that typically occur with only one copy per genome (V. Kunin *et al.*, unpublished data). To assess nucleotide sequence variation within the bins, we analysed the multiple alignment of the JAZZ assembly. Potential open reading frames (ORFs) were identified using 'mORFind' (J. Waldmann, H.T. and F.O.G., unpublished data) and annotation performed with the GenDB v2.2 system³⁶ and MicHanThi³⁷. The annotated symbiont metagenome was loaded into the metagenomics version of Integrated Microbial Genomes/M (IMG/M; <http://img.jgi.doe.gov/m>)³⁸.

Received 5 May; accepted 29 August 2006.

Published online 17 September; corrected 26 October 2006.

- Margulis, L. *Symbiosis in Cell Evolution* (W. H. Freeman, New York, 1993).
- Ruby, E. G., Henderson, B. & McFall-Ngai, M. We get by with a little help from our (little) friends. *Science* **303**, 1305–1307 (2004).
- DeLong, E. F. Microbial population genomics and ecology. *Curr. Opin. Microbiol.* **5**, 520–524 (2002).
- Handelsman, J. Metagenomics: application of genomics to uncultured microorganisms. *Microbiol. Mol. Biol. Rev.* **68**, 669–685 (2004).
- Riesenfeld, C. S., Schloss, P. D. & Handelsman, J. Metagenomics: genomic analysis of microbial communities. *Annu. Rev. Genet.* **38**, 525–552 (2004).
- DeLong, E. F. *et al.* Community genomics among stratified microbial assemblages in the ocean's interior. *Science* **311**, 496–503 (2006).
- Hallam, S. J. *et al.* Reverse methanogenesis: testing the hypothesis with environmental genomics. *Science* **305**, 1457–1462 (2004).
- Tyson, G. W. *et al.* Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* **428**, 37–43 (2004).
- Venter, J. C. *et al.* Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**, 66–74 (2004).
- Tringe, S. G. *et al.* Comparative metagenomics of microbial communities. *Science* **308**, 554–557 (2005).
- Wu, D. *et al.* Metabolic complementarity and genomics of the dual bacterial symbiosis of sharpshooters. *PLoS Biol.* **4**, e188 (2006).
- Dubilier, N., Blazejak, A. & Ruhland, C. Symbioses between bacteria and gutless marine oligochaetes. *Prog. Mol. Subcell. Biol.* **41**, 251–275 (2006).
- Blazejak, A., Erseus, C., Amann, R. & Dubilier, N. Coexistence of bacterial sulfide oxidizers, sulfate reducers, and spirochetes in a gutless worm (Oligochaeta) from the Peru margin. *Appl. Environ. Microbiol.* **71**, 1553–1561 (2005).
- Dubilier, N. *et al.* Endosymbiotic sulphate-reducing and sulphide-oxidizing bacteria in an oligochaete worm. *Nature* **411**, 298–302 (2001).
- Aparicio, S. *et al.* Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*. *Science* **297**, 1301–1310 (2002).
- Medini, D., Donati, C., Tettelin, H., Masignani, V. & Rappuoli, R. The microbial pan-genome. *Curr. Opin. Genet. Dev.* **15**, 589–594 (2005).
- Mowat, C. G. *et al.* Octaheme tetrathionate reductase is a respiratory enzyme with novel heme ligation. *Nature Struct. Mol. Biol.* **11**, 1023–1024 (2004).

18. van den Ende, F. P., Meier, J. & van Gemerden, H. Syntrophic growth of sulfate-reducing bacteria and colorless sulfur bacteria during oxygen limitation. *FEMS Microbiol. Ecol.* **23**, 65–80 (1997).
19. Matias, P. M., Pereira, I. A., Soares, C. M. & Carrondo, M. A. Sulphate respiration from hydrogen in *Desulfovibrio* bacteria: a structural biology overview. *Prog. Biophys. Mol. Biol.* **89**, 292–329 (2005).
20. Kletzin, A. & Adams, M. W. Molecular and phylogenetic characterization of pyruvate and 2-ketoisovalerate ferredoxin oxidoreductases from *Pyrococcus furiosus* and pyruvate ferredoxin oxidoreductase from *Thermotoga maritima*. *J. Bacteriol.* **178**, 248–257 (1996).
21. Cavanaugh, C. M., McKiness, Z. P., Newton, I. L. G. & Stewart, F. J. Marine chemosynthetic symbioses. In *The Prokaryotes: A Handbook on the Biology of Bacteria* (eds Dworkin, M. et al.) (Springer, New York, 2006).
22. Liu, J. Y., Miller, P. F., Gosink, M. & Olson, E. R. The identification of a new family of sugar efflux pumps in *Escherichia coli*. *Mol. Microbiol.* **31**, 1845–1851 (1999).
23. Giere, O. & Erseus, C. Taxonomy and new bacterial symbioses of gutless marine tubificidae (Annelida, Oligochaeta) from the island of Elba (Italy). *Org. Divers. Evol.* **2**, 289–297 (2002).
24. De Cian, M., Regnault, M. & Lallier, F. H. Nitrogen metabolites and related enzymatic activities in the body fluids and tissues of the hydrothermal vent tubeworm *Riftia pachyptila*. *J. Exp. Biol.* **203**, 2907–2920 (2000).
25. Khademi, S. et al. Mechanism of ammonia transport by Amt/MEP/Rh: structure of AmtB at 1.35 Å. *Science* **305**, 1587–1594 (2004).
26. Yancey, P. H., Blake, W. R. & Conley, J. Unusual organic osmolytes in deep-sea animals: adaptations to hydrostatic pressure and other perturbants. *Comp. Biochem. Physiol. A* **133**, 667–676 (2002).
27. Denger, K., Ruff, J., Schleheck, D. & Cook, A. M. *Rhodococcus opacus* expresses the xsc gene to utilize taurine as a carbon source or as a nitrogen source but not as a sulfur source. *Microbiology* **150**, 1859–1867 (2004).
28. Yang, C. C., Packman, L. C. & Scrutton, N. S. The primary structure of *Hyphomicrobium* X dimethylamine dehydrogenase. Relationship to trimethylamine dehydrogenase and implications for substrate recognition. *Eur. J. Biochem.* **232**, 264–271 (1995).
29. Paul, L., Ferguson, D. J. Jr & Krzycki, J. A. The trimethylamine methyltransferase gene and multiple dimethylamine methyltransferase genes of *Methanosarcina barkeri* contain in-frame and read-through amber codons. *J. Bacteriol.* **182**, 2520–2529 (2000).
30. Cohen, S. S. *A Guide to the Polyamines* (Oxford Univ. Press, New York, 1998).
31. Kelly, D. J. & Thomas, G. H. The tripartite ATP-independent periplasmic (TRAP) transporters of bacteria and archaea. *FEMS Microbiol. Rev.* **25**, 405–424 (2001).
32. Moran, N. A. Microbial minimalism: genome reduction in bacterial pathogens. *Cell* **108**, 583–586 (2002).
33. Moran, N. A. Tracing the evolution of gene loss in obligate bacterial symbionts. *Curr. Opin. Microbiol.* **6**, 512–518 (2003).
34. Moran, N. A. & Plague, G. R. Genomic changes following host restriction in bacteria. *Curr. Opin. Genet. Dev.* **14**, 627–633 (2004).
35. Huntemann, M. *MetaClust—Entwicklung eines modularen Programms zum Clustern von Metagenomfragmenten anhand verschiedener intrinsischer DNA-Signaturen*. Diploma thesis, Univ. Bremen (2006).
36. Meyer, F. et al. GenDB—an open source genome annotation system for prokaryote genomes. *Nucleic Acids Res.* **31**, 2187–2195 (2003).
37. Quast, C. *MicHanThi—Design and Implementation of a System for the Prediction of Gene Functions in Genome Annotation Projects*. Diploma thesis, Univ. Bremen (2006).
38. Markowitz, V. M. et al. The integrated microbial genomes (IMG) system. *Nucleic Acids Res.* **34**, D344–D348 (2006).
39. Teeling, H., Waldmann, J., Lombardot, T., Bauer, M. & Glockner, F. O. TETRA: a web-service and a stand-alone program for the analysis and comparison of tetranucleotide usage patterns in DNA sequences. *BMC Bioinformatics* **5**, 163 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was performed under the auspices of the US Department of Energy's Office of Science, Biological and Environmental Research Program, and supported by the University of California Lawrence Livermore National Laboratory, Lawrence Berkeley National Laboratory, and Los Alamos National Laboratory, and the Max Planck Society. We thank members of the Rubin, J. Bristow and P. Hugenholtz laboratories, as well as members of the Joint Genome Institute, for their contributions. We thank V. Markowitz, E. Dalin, N. Putnam, R. Sorek, T. Glavina del Rio, A. Salamov, A. Kobayashi and K. Kellaris for their assistance. At the Max Planck Institute for Marine Microbiology we thank S. Wetzel for technical assistance. We are grateful to C. Lott and the staff of the HYDRA field station at Elba for their generous support and help in sampling the worms.

Author Information The assembled sequences from the *Olavius* symbionts' metagenome have been deposited into the NCBI database under the project accession number AASZ00000000. The annotated *Olavius* symbionts' bins were incorporated into the metagenomics version of the US Department of Energy Joint Genome Institute Integrated Microbial Genomes/M (IMG/M; <http://img.jgi.doe.gov/m>). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to N.D. (ndubilie@mpi-bremen.de) or E.M.R. (EMRubin@lbl.gov).

ARTICLES

Complete crystallographic analysis of the dynamics of CCA sequence addition

Kozo Tomita¹, Ryuichiro Ishitani², Shuya Fukai² & Osamu Nureki^{2,3}

CCA-adding polymerase matures the essential 3'-CCA terminus of transfer RNA without any nucleic-acid template. However, it remains unclear how the correct nucleotide triphosphate is selected in each reaction step and how the polymerization is driven by the protein and RNA dynamics. Here we present complete sequential snapshots of six complex structures of CCA-adding enzyme and four distinct RNA substrates with and without CTP (cytosine triphosphate) or ATP (adenosine triphosphate). The CCA-lacking RNA stem extends by one base pair to force the discriminator nucleoside into the active-site pocket, and then tracks back after incorporation of the first cytosine monophosphate (CMP). Accommodation of the second CTP clamps the catalytic cleft, inducing a reorientation of the β -turn, which flips C74 to allow CMP to be accepted. In contrast, after the second CMP is added, the polymerase and RNA primer are locked in the closed state, which directs the subsequent A addition. Between the CTP- and ATP-binding stages, the side-chain conformation of Arg 224 changes markedly; this is controlled by the global motion of the enzyme and position of the primer terminus, and is likely to achieve the CTP/ATP discrimination, depending on the polymerization stage. Throughout the CCA-adding reaction, the enzyme tail domain firmly anchors the T Ψ C-loop of the tRNA, which ensures accurate polymerization and termination.

Genetic information is replicated and transcribed by polynucleotide polymerases. Conventional template-dependent DNA/RNA polymerases undergo transitions from 'open' to 'closed' conformations during the nucleotidyl transfer reaction. In the template-dependent DNA polymerases, nucleotide substrates are selected in the catalytically active closed conformation¹. In template-dependent RNA polymerases, nucleotide selection occurs in the catalytically inactive open conformation (pre-insertion state), whereas nucleoside incorporation occurs in the catalytically active closed conformation^{2,3}. In both cases, the template-dependent polymerases coordinate the dynamics of the enzyme, nucleic-acid template, primer and substrate.

The invariant CCA sequence found at the 3'-terminus of every tRNA is required for amino-acid acceptance and peptide-bond formation⁴⁻⁷. The CCA-adding RNA polymerase synthesizes and/or repairs the CCA sequence using CTP and ATP⁸, an activity that is conserved in all three primary kingdoms⁹. The CCA-adding enzyme catalyses three reactions without the aid of a nucleic-acid template: (1) CMP incorporation onto the 3'-termini of tRNAs that end with the discriminator (D) nucleotide (tRNA-D73); (2) CMP incorporation onto the 3'-termini of tRNAs that end with C74 (tRNA-DC74); and (3) AMP incorporation onto the 3'-termini of tRNAs that end with C75 (tRNA-DCC75). These reactions are expected to be accompanied by conformational transitions of the enzyme, RNA and substrate¹⁰. Biochemical studies indicate that CCA polymerization occurs in a single active pocket with the tRNA primer neither translocated nor rotated relative to the enzyme¹¹⁻¹³. Previous crystal structures have indicated that nucleotide specificity might be determined by hydrogen-bonding interactions that involve a single arginine residue and a phosphate group of the tRNA primer, with the size and shape of the nucleotide-binding pocket discriminating between CTP and ATP¹⁴. However, the primer was an RNA duplex that mimics the tRNA acceptor helix, and only the C75 and A76 polymerization states were visible. This RNA duplex lacked the T Ψ C-loop of tRNA, which

makes key interactions with the tail domain of the CCA-adding enzyme^{14,15}. Therefore, the precise mechanism by which the enzyme discriminates CTP from ATP as appropriate for the polymerization stage remains unclear.

To clarify the reaction dynamics of the CCA-adding enzyme, we solved the following six structures with a T Ψ C-loop-containing, 32-nucleotide CCA domain of tRNA (mini-helix): tRNA^{mini}-D73, tRNA^{mini}-DC74, tRNA^{mini}-DC74 + CTP, tRNA^{mini}-DCC75, tRNA^{mini}-DCC75 + ATP and tRNA^{mini}-DCCA76. These structures revealed the dynamic mechanism of nucleotide selection that accompanies CCA addition (see Supplementary Fig. 1), permitting us to compare the polymerization and termination dynamics of the template-dependent and template-free polynucleotide polymerases.

Overall structure

We solved crystal structures (2.5–2.8 Å) of the CCA-adding enzyme of *Archaeoglobus fulgidus* (AfCCA) complexed with four RNA primers (mini-helices) and incoming nucleotides, referred to as mini-D, mini-DC, mini-DC + CTP, mini-DCC, mini-DCC + ATP and mini-DCCA, covering primer binding, nucleotide selection, polymerization and termination. Apo-AfCCA (that is, tRNA-free AfCCA) forms a stable homodimer^{15,16}, with a reported enzyme:RNA primer stoichiometry of 2:2 (ref. 14). In our structures, one enzyme binds one RNA primer (Fig. 1a). AfCCA contains the head, neck, body and tail domains, as in previous studies¹⁴ (Fig. 1a). The tail recognizes the T Ψ C-loop of the mini-helix primer (Fig. 1b), and the 3'-terminus is located in the catalytic cleft formed by the head and neck (Fig. 1c). In the catalytic cleft, various primer nucleosides and incoming nucleotides move in and out of the relevant pockets, depending on the reaction stages. For clarity, we define the positions of the incoming nucleoside and tri-phosphate moieties, the priming position, and the two following stacking positions as the N, T, P, S1 and S2 sites, respectively (Fig. 1c).

¹Institute for Biological Resources and Functions, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1, Higashi, Tsukuba-shi, Ibaragi 305-8565, Japan. ²Department of Biological Information, Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Yokohama, Kanagawa 225-8501, Japan. ³SORST, JST, Honcho, Kawaguchi-shi, Saitama 332-0012, Japan.

Expansion and contraction of primer RNA

Figure 1b shows the interactions between the T Ψ C-loop and the tail domain in the mini-DCCA stage. Pro 347 in the tail domain stacks with the base moiety of C56 in the T Ψ C sequence (nucleosides 54, 55 and 56), which continuously stacks with G57. The phosphate groups of U(Ψ)55, C56 and G57 are recognized by the basic residues Arg 344, Lys 357, Arg 361 and Arg 363 (Fig. 1b). Asn 354 hydrogen bonds with the ribose O2' of C56 (Fig. 1b). On the other hand, interactions between the RNA acceptor/T Ψ C helices and the enzyme neck/body domains are maintained in all stages except mini-D (Fig. 1d). As reported previously¹⁴, these interactions include base-non-specific hydrogen-bonding and electrostatic interactions with the sugar-phosphate backbone of the tRNA (Fig. 1d).

In the mini-D stage, the acceptor helix is unexpectedly shifted by one nucleotide towards the catalytic cleft (+1 shift), owing to the RNA stem being stretched at G50 and C65 following disruption of the G50•C64 and G49•C65 base pairs (Fig. 2a and Supplementary Fig. 2). As a result, the extended regions comprising G7•C66, G49•C65, G50•C64 and C51•G63 deviate from the standard A-form helical conformation, with poor electron densities for G49 and G50 (Fig. 2a), indicating that entropy enhancement by their structural flexibilities might compensate for the enthalpy loss that results from the lack of base pairing. These disrupted base pairs are located beside the junction between the acceptor and T Ψ C-stems, exposed to the solvent and not directly recognized by any residue of the enzyme. The same amino-acid residues in the body and neck domains recognize the phosphate groups of the +1-shifted nucleoside, as compared to the other five stages, and the nucleosides in the T Ψ C-arm are recognized by the tail domain in the same manner (Fig. 2a, b). As a consequence, this 'expansion' of the primer RNA helix in the mini-D stage positions A73 and C72 at the S1 and S2 sites, respectively (Fig. 3a), so that A73 can accept CMP in the active-site pocket if an incoming CTP is accommodated. The electron densities of G1 and A73 were clearly visible in the unbiased $F_o - F_c$ omit map (Fig. 3a). After the first CMP

incorporation, the RNA helix snaps back to place A73 and the newly-formed C74 at the S2 and S1 sites, respectively (Figs 2b, 3b). Molecular simulation analyses of the internal motion of yeast tRNA^{Phe} by normal mode calculation found that the anticodon and acceptor stem regions thermally fluctuate independently of the residual regions of the RNA molecule^{17,18}, indicating that the tRNA junction region has some intrinsic plasticity. This suggests that the acceptor stem of the modified full-length tRNA also expands and contracts in the first step of the CCA-adding reaction.

CMP addition driven by 'knock-in' dynamics

In the mini-DC stage, the base moiety of C74 at the S1 site is sandwiched between A73 and the side chain of Glu 96 (Fig. 3b). The hydrogen-bond platforms formed between O ϵ 2 of Glu 96 and the main-chain amide group of Ala 126, and between N ϵ 2 of His 97 and O γ of Ser 125 might stabilize the stacking interaction between C74 and A73 (Fig. 3b and Supplementary Fig. 3a). In addition, O δ 1 of Asp 291 and N η 2 of Arg 224 form hydrogen bonds with the O2' of A73 and O3' of C74, respectively (Fig. 3b). Thus, the discriminator is recognized in a non-base-specific manner, and any base can be tolerated at position 73. The manner of primer recognition at the S1 and S2 sites is almost the same as that in the mini-D stage, except that the primer RNA is shifted by one nucleoside (Fig. 3a, b). The P and N sites are empty, and the O3' of C74 is distant from the three catalytic carboxylates of Glu 59, Asp 61 and Asp 110 in the mini-DC stage, representing a catalytically inactive state (Fig. 3b).

When CTP binds the binary complex (mini-DC + CTP stage), the head approaches the neck and body to adopt a closed conformation, in which α B, α C, and the loops connecting α B and β 1, β 2 and α C, and β 3 and β 4 relocate towards the neck domain (Figs 3c, 4). The largest conformational change is seen for the β -turn motif (residues 90–99) that connects β 3 and β 4 in the head domain; this change re-orientates Glu 96 and His 97, which were stacked onto C74 in the previous mini-DC stage (Figs 3c, 4). The terminal C74 of the primer is flipped and enters the P site (Figs 3c, 4); mutation of His 97 to Ala accelerated CMP addition onto tRNA^{mini-DC} (Supplementary Fig. 4a), supporting the functional importance of C74 flipping. In the P site, N4 and O2 of C74 form hydrogen bonds with the phosphate of A73 and the O γ 1 of Thr 130, respectively, with the 3'-OH group of C74 repositioned in the vicinity of the three catalytic carboxylates as well as the α -phosphate of incoming CTP. This represents a catalytically active state (Fig. 3c and Supplementary Fig. 3b). Incoming CTP is accommodated in the catalytic cleft constituted by the head and neck domains, as observed in previously reported structures¹⁴. A striking

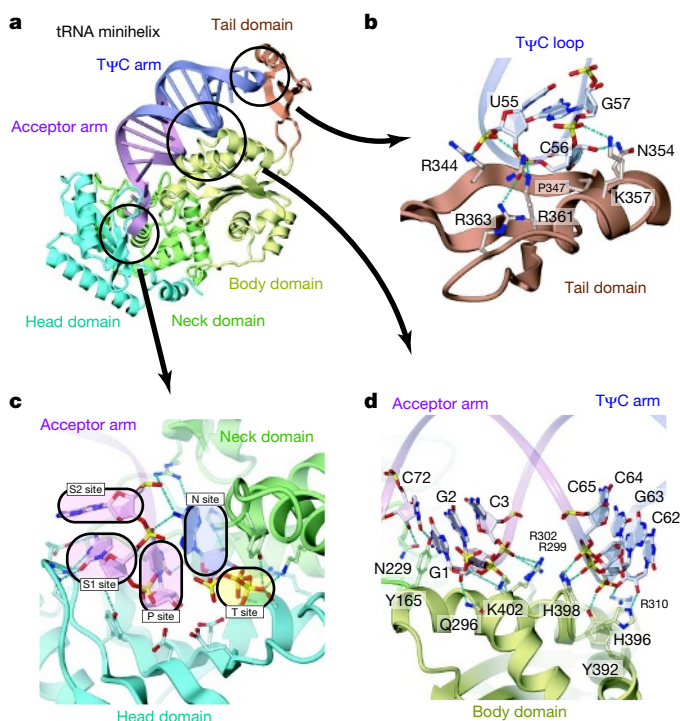


Figure 1 | Structure of the class I CCA-adding polymerase and tRNA mini-helix complex. **a**, Overall structure. **b**, Recognition of the tRNA T Ψ C-loop by the tail domain. **c**, Recognition of 3'-terminus of RNA at the catalytic site. Nomenclature of the nucleoside positions in the catalytic cleft is shown. **d**, Recognition of the tRNA acceptor/T Ψ C helices by the enzyme neck (green) and body (yellow) domains.

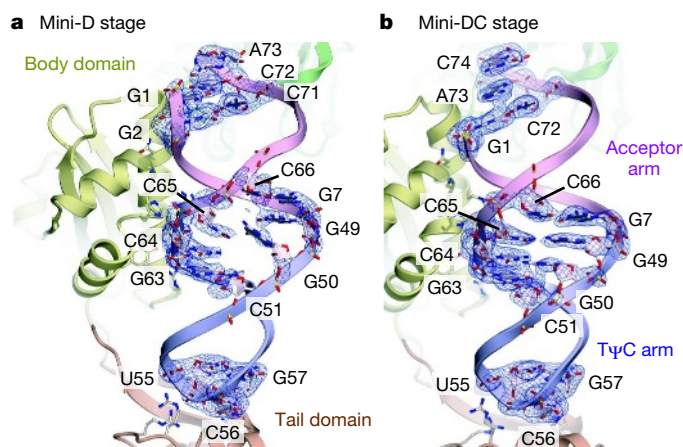


Figure 2 | Expansion and contraction of the primer RNA helix at the mini-D stage. **a**, Extended mini-helix structure at the mini-D stage. The simulated annealed omit maps (contoured at 3.5 σ) for the indicated nucleosides are shown. **b**, Back-tracked mini-helix structure at the mini-DC. Other stages have the same standard helix structure.

difference from previous structures is the side-chain conformation of Arg 224: its guanidinium group stretches towards the N site, owing to the approaching C74, to provide bipartite hydrogen bonds with the O2 and N3 of CTP (Fig. 3c), thereby eliminating the larger ATP. In the other stages, this side chain is bent to interact with the phosphate group of A73 (Fig. 3). In the previous, lower-resolution structure¹⁴,

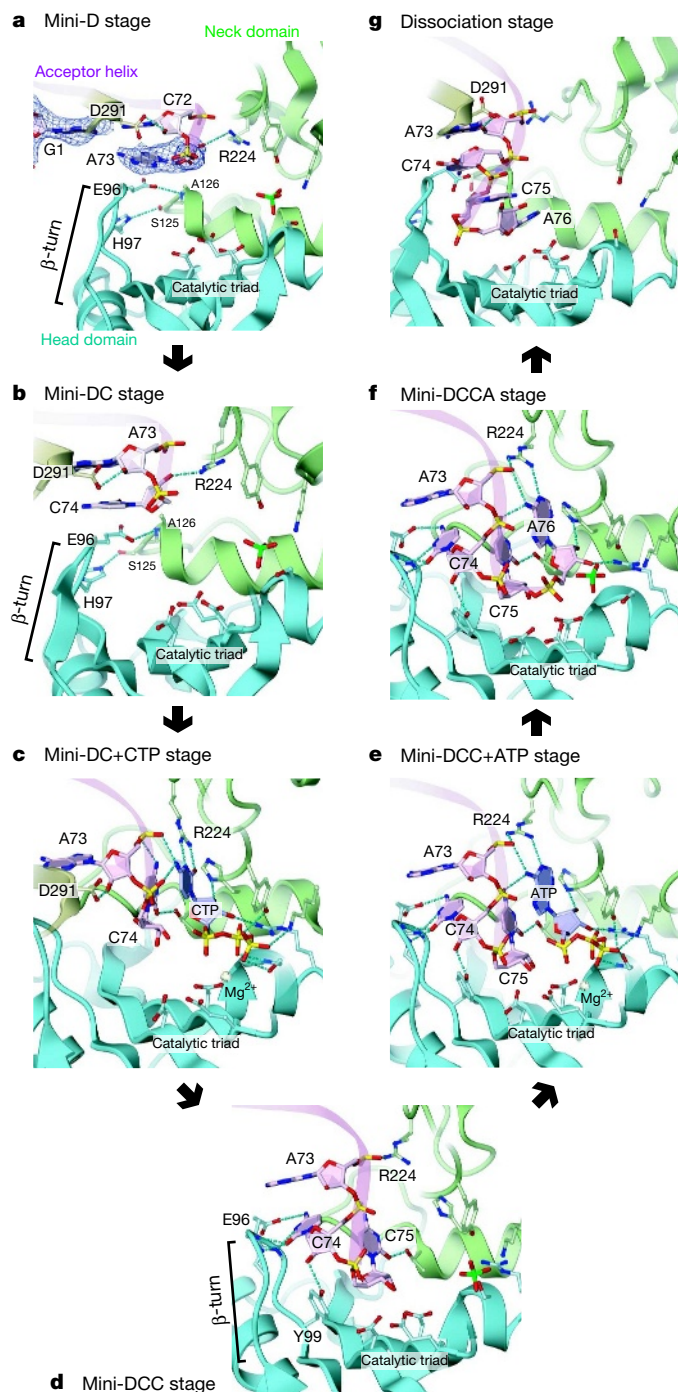


Figure 3 | Active-site structure at each reaction stage. **a**, Mini-D stage (the simulated annealed omit maps contoured at 4σ for G1 and A73 are shown). **b**, Mini-DC stage. **c**, Mini-DC + CTP stage. **d**, Mini-DCC stage. **e**, Mini-DCC + ATP stage. **f**, Mini-DCCA stage. **g**, Mature tRNA dissociation stage (PDB ID: 1SZ1)¹⁴. The acceptor stem expansion at the first CCA-adding step can easily be seen by the fact that Asp 291 recognizes the 2'-OH group of C72 in the mini-D stage, but A73 in the other stages. The catalytic triad comprises Glu 59, Asp 61 and Asp 110. The hydrogen bonds are represented by dotted lines.

the side-chain conformation of Arg 224 appeared the same for the CTP- and ATP-adding stages.

On the basis of these observations, we propose a new CTP selection mechanism, in which three events facilitate transition from the mini-DC to the mini-DC + CTP stages: (1) correct interactions between the sugar/triphosphate moieties of an NTP and the polymerase head/neck domains; (2) an open to closed conformational transition, which causes structural change of the β -turn motif and C74 flipping (Fig. 4); and (3) proper re-packing of the side chain of Arg 224. Only CTP can properly induce all three events, so it is selectively incorporated into the catalytic pocket. ATP produced no clear electron density when soaked into the mini-DC-stage crystal (data not shown). These results indicate that nucleotide-induced 'knock-in' dynamics must be the underlying mechanism for the precise selection of CTP (Fig. 4). Consistent with these structural observations, replacements of Arg 224, Arg 50, Ser 47 and Tyr 161 with Ala all reduced CMP incorporation (Supplementary Fig. 4a).

AMP addition in the locked catalytic pocket

After the second CMP incorporation at position 75, C74 flips back to the S1 site. The newly formed C75 shifts to the P site, where its 3'-OH group locates close to the catalytic triad (Fig. 3d and Supplementary Fig. 3c). In this mini-DCC stage, although C74 re-occupies the same S1 site as in the mini-DC stage, the binding mode differs from that of the mini-DC stage (Fig. 3b, d). The base moiety of C74 is recognized by Watson-Crick type hydrogen bonds with the β -turn motif (Glu 96 and His 97; Fig. 3d), which pinned down C74 by stacking interactions in the preceding mini-DC stage (Fig. 3b). The 4-amino group of C75 hydrogen bonds with the phosphate group of C74, while O2 of C75 forms a hydrogen bond with the β -hydroxyl group of Thr 130 (Fig. 3d). Intriguingly, the entire polymerase conformation in the mini-DCC stage remains locked in the active closed state, despite the lack of an NTP, which can be ascribed to the Watson-Crick type 'locking' interactions between the β -turn motif and C74 at the S1 site (Fig. 3d). The Arg 224 side chain now returns to the bent conformation.

When ATP binds the binary complex (mini-DCC + ATP stage), the enzyme maintains a closed state without any significant conformational change of the primer 3'-end or the progressive repositioning of the enzyme head domain, in contrast to previous reports¹⁴ (Fig. 3e and Supplementary Fig. 5). Consequently, the Arg 224 side

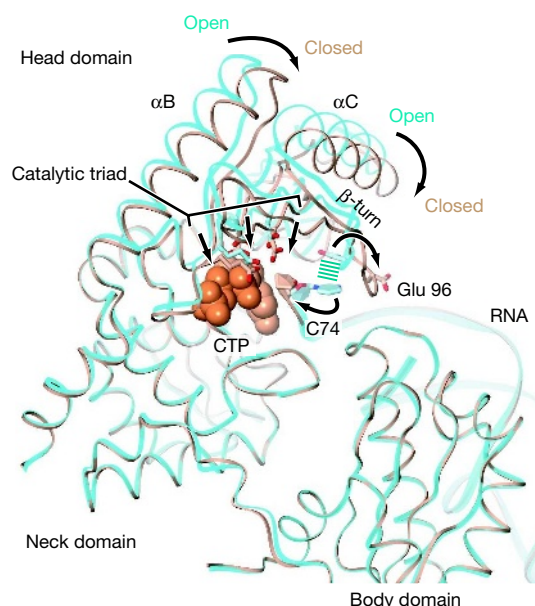


Figure 4 | Open-to-closed transition upon the accommodation of CTP at the mini-DC + CTP stage. During the conformational transition, Glu 96 relocates to allow the terminal C74 to flip.

chain remains bent to hydrogen bond with the A73 phosphate group (Fig. 3e), and the N site can then accommodate the larger ATP (Fig. 3e). Incoming ATP binds the N/T pocket (Fig. 3e and Supplementary Fig. 6) that was formed in the previous mini-DCC stage (Fig. 3d), indicating that binding of tRNA^{mini}-DCC to the enzyme might be sufficient for ATP selection. This differs from CMP incorporation, where incoming CTP is selected by the dynamic open to closed transition involving Arg 224 reorientation (Figs 3c, 4). When we soaked CTP into the mini-DCC stage, Arg 224 adopted the same bent conformation as the mini-DCC + ATP stage (Supplementary Fig. 7), indicating that the side-chain conformation of Arg 224 does not simply fit CTP but is controlled by the enzyme's global motion and primer terminus repositioning. When we used Ala to replace Arg 50, Ser 47 and Tyr 161 (all of which interact with the triphosphate moiety) and Thr 130 (which interacts with the 2'-OH group of C75; Fig. 3e), AMP incorporation activity was reduced, as expected (Supplementary Fig. 4b). In contrast, the R224A variant, which reduced CMP addition onto tRNA^{mini}-DC, did not impair AMP incorporation (Supplementary Fig. 4b). These results show that the main determinant of precise ATP selection is the static active-site structure formed by the polymerase and 3'-terminus of the RNA primer locked in the closed conformation.

Termination by pyrophosphate dissociation

In the mature mini-DCCA stage, there is no structural change of the enzyme, which remains locked in the closed conformation, and the newly added 3'-terminal A76 is still locked in the N site (Fig. 3f and Supplementary Fig. 3e). However, the conformation of the primer 3'-terminus in the mini-DCCA stage differs from the previously reported complex structure with mature tRNA¹⁴ (Fig. 3g). In the previously reported structure, the polymerase adopted a more open conformation than at any other stage, and C75 and A76 escaped from the P and N sites, respectively. C74, C75 and A76 were stacked on each other; C74 remained at the S1 site but lacked Watson-Crick type 'locking' interactions with the β -turn motif (Fig. 3g). In the mini-DCCA stage, we observed electron density for a sulphate ion close to the γ -phosphate of the incoming nucleotide at the T site (Fig. 3f and Supplementary Fig. 3e): this sulphate ion accepts hydrogen bonds from Lys 152, Tyr 161 and Ser 47 as well as the NTP γ -phosphate. It is reasonable to assume that the sulphate mimics the phosphate moiety of pyrophosphate (PPi), in which case the mini-DCCA structure would represent the reaction status immediately after nucleotidyl transfer, with PPi retained. The previously reported complex structure with mature tRNA¹⁴ is assumed to be the termination stage with PPi already released. PPi dissociation might reopen the catalytic cleft, leading to a conformational change of the primer 3'-terminus and thereby releasing the mature tRNA.

Vice-anchored knock-in-and-lock dynamics

Previous structures of the AfCCA complex with RNA duplexes indicated that ligand selectivity might be achieved through a progressive opening up of the flexible head domain so that the size and shape of the nucleotide-binding pocket change^{14,19}. This is in striking contrast to our results, which show that the mini-DC+CTP and mini-DCC + ATP structures are almost the same, except for the relocation of Arg 224 (Supplementary Fig. 5). This difference might be ascribed to the primer that was used or to the resolution of the structure determination. Our sequential structures provide a clear picture, showing that the CCA-adding polymerase adopts two discrete conformations: the open conformation for the mini-D and mini-DC stages, and the closed conformation for the mini-DC + CTP, mini-DCC, mini-DCC + ATP and mini-DCCA stages (Supplementary Fig. 8). CTP accommodation triggers this open-to-closed transition, inducing a structural change of the β -turn motif that allows the terminal nucleoside of the primer to flip and Arg 224 to relocate (Supplementary Fig. 8). This mutual adaptation between CTP, the RNA primer and the polymerase specifies CTP and knocks it onto the

primer terminus. Thus, C addition is driven by the 'knocking-in' dynamics of the polymerase and primer (Supplementary Movies 1 and 2). The open-to-closed transition that is triggered by the incoming NTP is quite similar to that observed in conventional DNA polymerases¹. After the second CMP addition, the conformation of the polymerase is fixed in the closed state, by 'locking' interactions between C74 and the β -turn motif, and the static, locked structure of the NTP-binding pocket directs the addition of ATP (Supplementary Movies 1 and 2). This is in striking contrast to the continuous polymerization that is catalysed by conventional template-dependent DNA/RNA polymerases¹⁻³. The tight anchoring of the tip of the acceptor-T Ψ C helix ensures that polymerization terminates after CCA is completed. As with the conventional polymerases³, dissociation of the pyrophosphate by-product drives the polymerization cycle to the next step, allowing the catalytic cleft to reopen and release the mature tRNA.

METHODS

Crystallization and structural determination. AfCCA was prepared as described¹⁵. Mini-helices derived from *Thermotoga maritima* tRNA^{Phe} were synthesized by T7 RNA polymerase *in vitro* and purified by polyacrylamide gel electrophoresis. The purity of the mini-helices was >99%, as judged from electrophoresis and *in vitro* CCA-adding reactions. Crystallization and data collection of the AfCCA binary and ternary complexes are described in the Supplementary Information. All of the data were processed using the HKL2000 program²⁰. Crystal structures were solved at 2.5–2.8 Å resolutions by molecular replacement with the program AMORE²¹, using the refined Apo-AfCCA structure¹⁵ (Protein Data Bank ID: 1UET) as a search model. The model was manually modified using the program O (ref. 22), with iterative cycles of refinement with CNS (ref. 23) (Supplementary Table 1). To verify the divalent cation binding at the catalytic site, we soaked the mini-DCC + ATP ternary complex in 5 mM MnCl₂. A data set was collected at a wavelength of 1.5 Å, and anomalous difference Fourier synthesis was performed (Supplementary Fig. 6). All molecular figures were created using the program CueMol (<http://www.cuemol.org/>).

***In vitro* CCA-adding assay.** Mutations were introduced into the AfCCA-over-expressing plasmid by site-directed mutagenesis, and the mutant enzymes were prepared as described previously¹⁵. The incorporation of AMP into tRNA^{mini}-DCC75 was measured in a buffer containing 50 mM glycine-NaOH pH 8.5, 10 mM MgCl₂, 2 mM DTT, 30 μ M ATP, 82.5 nM [α -³²P]ATP (3,000 Ci mmol⁻¹), 5 μ M tRNA^{mini}-DCC75 and 50 ng enzyme at 50 °C for 10 min. For the CMP incorporation assay, 30 μ M CTP, 82.5 nM [α -³²P]CTP and 5 μ M tRNA^{mini}-DC74 were used. The K_m values for ATP and CTP were estimated as 283 and 56 μ M, respectively, and the K_m values for tRNA^{mini}-DCC75 and tRNA^{mini}-DC74 were estimated as 1.14 and 0.99 μ M, respectively. We quantified the nucleotide incorporation rate as described^{10,15}.

Received 8 June; accepted 6 September 2006.

Published online 15 October 2006.

1. Doublié, S. & Ellenberger, T. The mechanism of action of T7 DNA polymerase. *Curr. Opin. Struct. Biol.* **8**, 704–712 (1998).
2. Temiakov, D. *et al.* Structural basis for substrate selection by T7 RNA polymerase. *Cell* **116**, 381–391 (2004).
3. Yin, Y. W. & Steitz, T. A. The structural mechanism of translocation and helicase activity in T7 RNA polymerase. *Cell* **116**, 393–404 (2004).
4. Sprinzl, M. & Cramer, F. The -C-C-A end of tRNA and its role in protein biosynthesis. *Prog. Nucleic Acid Res. Mol. Biol.* **22**, 1–69 (1979).
5. Green, R. & Noller, H. F. Ribosomes and translation. *Annu. Rev. Biochem.* **66**, 679–716 (1997).
6. Kim, D. F. & Green, R. Base-pairing between 23S rRNA and tRNA in the ribosomal A site. *Mol. Cell* **4**, 859–864 (1999).
7. Nissen, P., Hansen, J., Ban, N., Moore, P. B. & Steitz, T. A. The structural basis of ribosome activity in peptide bond synthesis. *Science* **289**, 920–930 (2000).
8. Deutscher, M. P. in *Enzymes of Nucleic Acid Synthesis and Modification*. RNA Enzymes vol. 2 (ed. Jacob, S. T.) 159–183 (CRC, Boca Raton, Florida, 1983).
9. Yue, D., Maizels, N. & Weiner, A. M. CCA-adding enzymes and poly(A) polymerases are all members of the same nucleotidyltransferase superfamily: characterization of the CCA-adding enzyme from the archaeal hyperthermophile *Sulfolobus shibatae*. *RNA* **2**, 895–908 (1996).
10. Tomita, K. *et al.* Structural basis for template-independent RNA polymerization. *Nature* **430**, 700–704 (2004).
11. Weiner, A. M. tRNA maturation: RNA polymerization without a nucleic acid template. *Curr. Biol.* **14**, 883–885 (2004).
12. Yue, D., Weiner, A. M. & Maizels, N. The CCA-adding enzyme has a single active site. *J. Biol. Chem.* **273**, 29693–29700 (1998).

13. Shi, P. Y., Maizels, N. & Weiner, A. M. CCA addition by tRNA nucleotidyltransferase: polymerization without translocation? *EMBO J.* **17**, 3197–3206 (1998).
14. Xiong, Y. & Steitz, T. A. Mechanism of transfer RNA maturation by CCA-adding enzyme without using an oligonucleotide template. *Nature* **430**, 640–645 (2004).
15. Okabe, M. *et al.* Divergent evolutions of trinucleotide polymerization revealed by an archaeal CCA-adding enzyme structure. *EMBO J.* **22**, 5918–5927 (2003).
16. Xiong, Y., Li, F., Wang, J., Weiner, A. M. & Steitz, T. A. Crystal structures of an archaeal class I CCA-adding enzyme and its nucleotide complexes. *Mol. Cell* **12**, 1165–1172 (2003).
17. Nakamura, S. & Doi, J. Dynamics of transfer RNAs analyzed by normal mode calculation. *Nucleic Acids Res.* **22**, 514–521 (1994).
18. Matsumoto, A., Tomimoto, M. & Go, N. Dynamical structure of transfer RNA studied by normal mode analysis. *Eur. Biophys. J.* **28**, 369–379 (1999).
19. Xiong, Y. & Steitz, T. A. A story with a good ending: tRNA 3'-end maturation by CCA-adding enzymes. *Curr. Opin. Struct. Biol.* **16**, 1–6 (2006).
20. Otwinowski, Z. & Minor, W. in *Methods in Enzymology* vol. 276 (eds Carter, C. W. & Sweet, R. M.) 307–326 (Academic, New York, 1997).
21. Navaza, J. Amore: an automated package for molecular replacement. *Acta Crystallogr. A* **50**, 157–163 (1994).
22. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
23. Brunger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Ibba for improvement of the manuscript. We thank the beamline staff at BL41XU of SPring-8 and BL5 and NW-12 of KEK for technical help during data collection, and A. Hamada for technical assistance. This work was supported by grants from JSPS for young scientists, MEXT for priority areas of science, the Mitsubishi Foundation, the Sumitomo Foundation and AIST to K.T., as well as by a PRESTO and SORST Program grant from JST (Japan Science and Technology) and a grant from MEXT to O.N.

Author Contributions K.T. carried out structural determination and mutant analysis, and wrote the paper with editing from R.I. and O.N. R.I., S.F. and O.N. assisted with the structural determination. All authors discussed the results and commented on the manuscript. O.N. supervised the work.

Author Information Coordinates and structure factors have been deposited in the Protein Data Bank, under the accession codes 2DR5, 2DR7, 2DR8, 2DR9, 2DRA, 2DRB and 2DVI for the mini-D stage, the mini-DC stage, the mini-DC + CTP stage, the mini-DCC stage, the mini-DCC + ATP stage, the mini-DCCA stage and the mini-DCC + CTP misrecognition complex, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to K.T. (kozo-tomita@aist.go.jp) or O.N. (nureki@bio.titech.ac.jp).

Evidence for superfluidity of ultracold fermions in an optical lattice

J. K. Chin¹, D. E. Miller¹, Y. Liu¹, C. Stan¹†, W. Setiawan¹, C. Sanner¹, K. Xu¹ & W. Ketterle¹

The study of superfluid fermion pairs in a periodic potential has important ramifications for understanding superconductivity in crystalline materials. By using cold atomic gases, various models of condensed matter can be studied in a highly controllable environment. Weakly repulsive fermions in an optical lattice could undergo *d*-wave pairing¹ at low temperatures, a possible mechanism for high temperature superconductivity in the copper oxides². The lattice potential could also strongly increase the critical temperature for *s*-wave superfluidity. Recent experimental advances in bulk atomic gases include the observation of fermion-pair condensates and high-temperature superfluidity^{3–8}. Experiments with fermions^{9–11} and bosonic bound pairs^{12,13} in optical lattices have been reported but have not yet addressed superfluid behaviour. Here we report the observation of distinct interference peaks when a condensate of fermionic atom pairs is released from an optical lattice, implying long-range order (a property of a superfluid). Conceptually, this means that *s*-wave pairing and coherence of fermion pairs have now been established in a lattice potential, in which the transport of atoms occurs by quantum mechanical tunnelling and not by simple propagation. These observations were made for interactions on both sides of a Feshbach resonance. For larger lattice depths, the coherence was lost in a reversible manner, possibly as a result of a transition from superfluid to insulator. Such strongly interacting fermions in an optical lattice can be used to study a new class of hamiltonians with interband and atom–molecule couplings¹⁴.

Previous experiments showing long-range phase coherence in Bose–Einstein condensates (BECs) and in fermion superfluids used ballistic expansion to observe the interference of two independent condensates¹⁵, vortex lattices^{8,16,17} or interference peaks after release from an optical lattice^{18,19}. However, for strongly interacting fermions, elastic collisions can change the momentum distribution and wash out interference peaks. For an initially superfluid cloud, such dissipative dynamics corresponds to superfluid flow faster than the critical velocity. Consistent with this expectation is the observation that a strongly interacting Fermi superfluid initially containing distinct momentum components yielded a broad diffuse cloud after expansion (Fig. 1). This issue was addressed by using a magnetic field ramp that quickly increased the detuning from a Feshbach resonance, taking the system out of the strongly interacting regime and enforcing ballistic expansion. In previous studies of strongly interacting Fermi gases, magnetic field sweeps were applied to prevent fermion pairs above the Feshbach resonance from dissociating^{6,7,20}. In contrast, our experiment required a magnetic field sweep both above and below the Feshbach resonance to avoid elastic collisions.

Our experiments used a balanced mixture of ⁶Li fermions in the two lowest hyperfine states. Evaporative cooling produced a nearly pure fermion pair condensate that was adiabatically loaded into a

three-dimensional optical lattice. A broad Feshbach resonance centred at 834 G enabled tuning of the interatomic interactions over a wide range. On resonance, a bound molecular state becomes degenerate with the open atomic scattering channel, leading to a divergence in the scattering length *a*. Here we explore the region of strong interactions, also known as the BEC–BCS (Bardeen–Cooper–Schrieffer) crossover, in which the magnitude of the interaction parameter $|k_F a|$ is greater than unity, and k_F is defined as the peak Fermi wavevector of a two-component non-interacting mixture of ⁶Li atoms. In the crossover region, pairing occurs as a result of many-body interactions. Below resonance, for strong interactions, the bare two-body state has a bond length larger than the interatomic spacing and is irrelevant. In a lattice, atom pairs above and below the resonance can be confined to one lattice site¹¹, and crossover physics may require an occupation larger than or equal to one.

The peak filling factor of the lattice was about unity. At this density in the bulk, the fermion pair size is on the order of $1/k_F = 170$ nm, comparable to the lattice spacing of 532 nm. To probe the momentum distribution, we ramped the magnetic field out of the strongly interacting regime as fast as technically possible (about 150 μ s) and then turned off the confining potential. Absorption

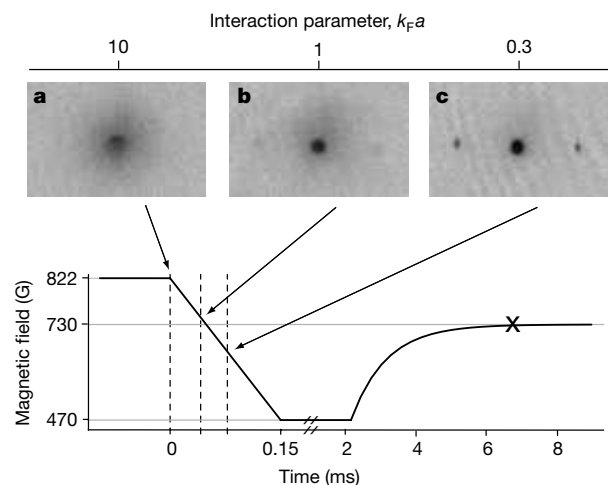


Figure 1 | Dissipative collisions during expansion of a strongly interacting fermionic superfluid. The schematic shows the time sequence of the magnetic field ramp used throughout this paper. A one-dimensional optical standing wave was pulsed onto the superfluid at different magnetic fields B_p (indicated by arrows at 822 G (a), 749 G (b) and 665 G (c)) during expansion, creating particles at twice the photon recoil³⁰. Absorption images taken at the time marked with the cross show distinct momentum peaks only at magnetic fields $B_p \leq 750$ G, where $k_F a \leq 1$. At higher magnetic fields, the peaks blurred into a broad diffuse cloud as a result of the larger collision cross-section.

¹Department of Physics, MIT–Harvard Center for Ultracold Atoms, and Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. †Present address: Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, USA.

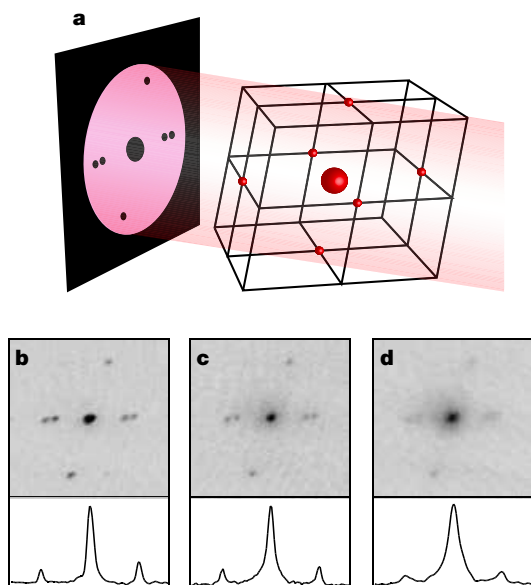


Figure 2 | Observation of high-contrast interference of fermion pairs released from an optical lattice below and above the Feshbach resonance. **a**, The orientation of the reciprocal lattice, also with respect to the imaging light. **b–d**, Interference peaks are observed for magnetic fields of 822 G (**b**), 867 G (**c**) and 917 G (**d**). The lattice depth for all images is $5E_r$, and each image is the average of three shots. The field of view is $1\text{ mm} \times 1\text{ mm}$. Density profiles through the vertical interference peaks are shown for each image.

images taken after 6.5 ms of expansion reveal sharp peaks at the reciprocal lattice vectors—the signature of long-range coherence, a strong indicator for superfluidity.

We observed these interference peaks at magnetic fields both above and below the Feshbach resonance (Fig. 2). The six first-order diffracted peaks are clearly visible around the zero momentum fraction and their positions correspond to the expected momentum quanta of $2\hbar k_L$ carried by molecules of mass $2m$, where k_L is the lattice wavevector. At high magnetic fields (Fig. 2d) the visibility of the interference peaks decreased and some additional heating was observed. This degradation could be due to a higher fraction of thermal atoms as we approached the BCS limit, but it was not studied in detail.

The narrow interference peaks clearly reveal the presence of a macroscopic wavefunction possessing long-range phase coherence. The separation between the interference peaks relative to their width gives an estimate of the coherence length of about ten lattice sites. This estimate is a lower bound, because effects of finite resolution and mechanisms of residual broadening have been neglected. With unity occupation, and in the absence of any discernible background at magnetic fields near the Feshbach resonance, this implies a minimum phase space density of 10^3 and shows that our samples are deep in the quantum-degenerate regime. In previous studies of ultracold Bose and Fermi gases, the appearance of a condensate fraction and long-range phase coherence was shown to occur concurrently with the possibility to excite superfluid flow^{8,16,17,21}. Superfluid hydrodynamics is usually regarded as the direct proof for superfluidity. However, all reports of superfluidity of bosons in three-dimensional optical lattices have relied solely on observations of sharp interference peaks and inferred superfluidity from the established connection between long-range coherence and superfluidity^{19,22}. Similarly, our observations directly show long-range coherence and indirectly show superfluidity of fermion pairs in an optical lattice.

For deep lattices, breakdown of superfluid behaviour has been observed for weakly interacting BECs of different bosonic species^{19,23}. This phase transition to the Mott-insulator state occurs when on-site interactions start to suppress atom number fluctuations and the system undergoes a transition from a delocalized superfluid

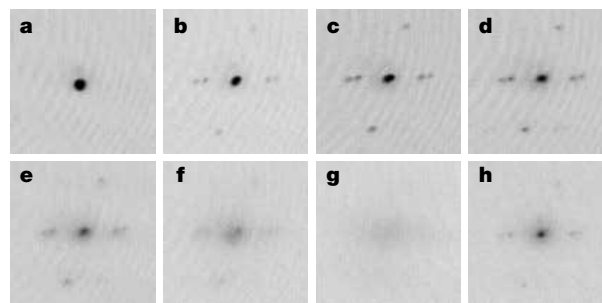


Figure 3 | Interferograms of fermion pairs released from different lattice depths V_0 at a field of 822 G. Values of V_0 are $0E_r$ (**a**), $2.5E_r$ (**b**), $4E_r$ (**c**), $5E_r$ (**d**), $6E_r$ (**e**), $7E_r$ (**f**), $9E_r$ (**g**) and $2.5E_r$ (**h**). **a–g** were taken after an adiabatic ramp up to the final V_0 , whereas **h** was taken after first ramping up to $10E_r$ before ramping down to $2.5E_r$.

described by a macroscopic wavefunction to a product of Wannier states tightly localized at each lattice site. Experimentally, this is manifested as a smearing of the distinct $2\hbar k_L$ interference peaks.

Figure 3 shows the change in the coherence properties when the lattice depth was increased. The interference peaks became more pronounced initially, because of increased modulation of the wavefunction. The interference peaks began to smear out, rapidly giving way to a featureless cloud, beyond a critical lattice depth $V_c \approx 6E_r$, where $E_r = \hbar^2 k_L^2 / 4m = h \times 15\text{ kHz}$ is the recoil energy. This indicates that all phase coherence had been lost. On subsequent ramping down of the lattice, interference peaks became visible again (Fig. 3h), showing reversibility of the lattice ramp.

We repeated this sequence for a wide range of initial magnetic fields, both above and below the resonance, and observed the same marked change in the interference pattern. Figure 4 displays the peak optical density of the interference peaks for different lattice depths at representative fields. Across all fields, the sharp decrease in peak optical density occurred between $5E_r$ and $6E_r$. A further increase in the magnetic field resulted in decreasing overall visibility, until interference peaks could no longer be observed regardless of lattice depth.

The loss of phase coherence with increasing lattice depth is consistent with the qualitative description of the superfluid to Mott-insulator transition. However, the usual single-band description is no longer applicable, because in the strong-coupling regime the on-site interaction strength should be comparable to the band gap $\hbar\omega$, where ω is the onsite trap frequency. Furthermore, Pauli blocking forbids the multiple occupation of the lowest state of an

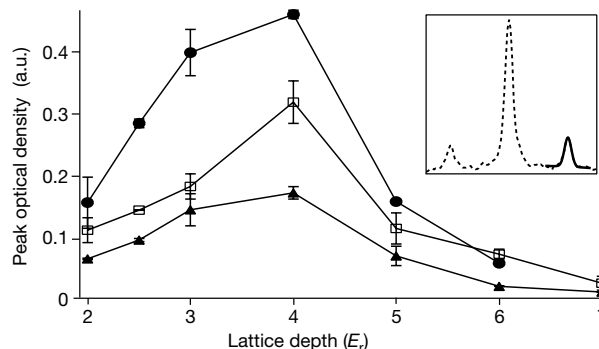


Figure 4 | Peak optical density of interference peaks for increasing lattice depths at different magnetic fields. Values of magnetic fields are 842 G (filled circles), 892 G (open squares) and 942 G (filled triangles). Peak optical densities were estimated from fits to the peaks, including background subtraction. The inset shows a sample density profile of the central and one pair of interference peaks (dotted line), with a bimodal fit to one side peak (solid line). Each point is the average of three different images with six interference peaks per image. Error bars show s.d.

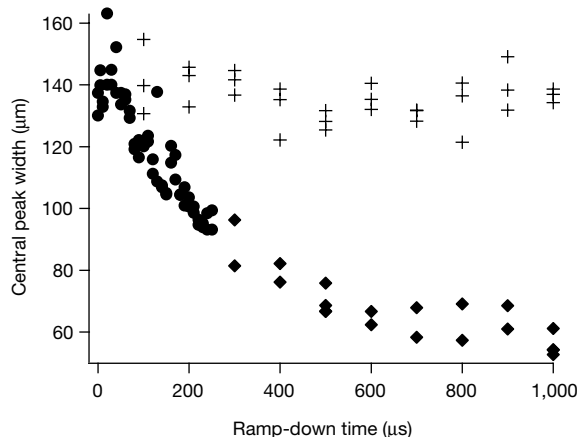


Figure 5 | Restoring coherence from a deep lattice. The width of the central peak is used as a measure of phase coherence after an adiabatic ramp up to $8E_F$, followed by a fast ramp down to $2.5E_F$ at a fixed magnetic field of 822 G. Filled circles were extracted with the use of a gaussian fit, and diamonds with a bimodal fit. Also plotted for comparison is the gaussian width of the central peak for a dephased sample, in which a field gradient was applied during the ramp up of the lattice (crosses). All points were taken for 6.5 ms time of flight.

individual lattice site by identical fermions, and modification of the single-particle tunnelling rate is expected as a result of virtual pair-breaking transitions¹⁴. One may still be tempted to use the standard bosonic Hubbard model and estimate the critical lattice depth V_c for an assumed value of onsite interaction energy $U = \hbar\omega$ and non-interacting, single-particle tunnelling J , but the obtained $V_c \approx 3E_F$ is significantly smaller than our observation, which is in turn much smaller than the $V_c > 10E_F$ observed for weakly interacting atomic BECs^{19,23}. Together with the observed insensitivity of V_c to the magnetic field, this shows that models based on weak interactions are inadequate.

Figure 3h shows the reversibility of the transition from a long-range coherent state to a state without strong coherence. We now study the timescale for this recoherence, by analogy with similar measurements performed across the transition from superfluid to Mott insulator in atomic BECs¹⁹. Figure 5 shows that phase coherence was restored on a submillisecond timescale, on the order of the single-particle tunnelling time of about 500 μ s (for a shallow lattice of $2.5E_F$). When the same lattice ramp sequence was applied to a superfluid that had been dephased by a magnetic field gradient¹⁹, the system did not regain phase coherence on the timescales that we probed. Evaporative cooling is therefore negligible during this time. The short recoherence time of the condensate is evidence that the system stayed in its ground state or at least in a low-entropy state when the lattice was ramped up.

Figure 5 also provides evidence that the system could not recohere during the 150- μ s magnetic field ramp. In Fig. 3h, the central peak is well fitted by a bimodal distribution with a width of 35 μ m, in clear contrast to the gaussian width of 105 μ m obtained from Fig. 5 after 150 μ s. We therefore conclude that the observed interference patterns in Fig. 1 reflect the coherence of the cloud at the initial magnetic field, in the strongly interacting regime.

We have shown long-range phase coherence of fermion pairs in an optical lattice in the BEC–BCS crossover region by observing sharp interference peaks during ballistic expansion. This indicates that we have achieved *s*-wave pairing and superfluidity in a lattice potential. Further studies will reveal how the pair wavefunction is affected by confinement²⁴, and whether the lattice shifts the BEC–BCS crossover away from the Feshbach resonance²⁵. The loss of coherence during the lattice ramp up and the rapid recoherence are characteristic of a Mott insulator. However, definitive proof will require a

better understanding of the unitarity-limited interactions in such a Fermi system. Recent theoretical work^{14,26} predicts that strongly interacting fermions in an optical lattice feature multiband couplings and next-neighbour interactions and can realize the important *t*–*J* and magnetic XXZ models of condensed-matter theory. This demonstrates that such atomic systems are an ideal laboratory for the exploration of novel condensed-matter physics.

METHODS

Clouds of superfluid fermion pairs were created in a new experimental setup^{27,28} by using techniques similar to those described elsewhere⁸. In brief, a combination of laser cooling and sympathetic cooling of spin-polarized fermions by bosonic ^{23}Na was followed by a spin transfer to create a two-component Fermi gas, allowing further cooling through direct evaporation of the fermions. As the fermions cooled, they formed pairs that Bose-condensed.

Estimates of the scattering length, and hence the interaction parameter, from the magnetic field were obtained with $a(B) = -1,405a_0[1 + 300/(B - 834)][1 + 0.0004(B - 834)]$ (ref. 29), where B is measured in gauss and a_0 is the Bohr radius. The calibration of the magnetic field in our system had an uncertainty of about 5 G.

Evaporation was performed at a magnetic field of 822 G, at which strong interactions permitted efficient evaporation. An estimated average final number of $N \approx 2 \times 10^5$ ^6Li pairs and harmonic trapping frequencies of $\nu_{x,y,z} = (270,340,200)$ Hz gave a trap depth of 1.7 μ K and a Fermi energy of $E_F = k_B \times 1.4$ μ K, where $E_F = \hbar\omega_F(6N)^{1/3}$ and ω_F is the average trapping frequency. After evaporation, the magnetic field was brought to a desired value B_0 in 20 ms and the condensate was allowed to equilibrate for a further 200 ms. Before ramping to values of B_0 on the BCS side, we also recompressed the optical trap to (340,440,270) Hz and 2.2 μ K depth in 100 ms to accommodate the larger Fermi clouds above the resonance⁷.

A three-dimensional optical lattice was formed from three optical standing waves, oriented such that the resulting unit cell had a sheared cubic structure, with one axis tilted about 20° from the normal for reasons of optical access (see Fig. 1a)²³. The incident laser beams were focused down to the condensate with waists of about 90 μ m, then retroreflected and overlapped at the condensate to generate the standing-wave potentials. All lattice light was derived from a 1,064-nm single-frequency fibre laser, and each beam was detuned by tens of MHz with respect to the others to eliminate interference between different beams.

The lattice potential was imposed on the condensate by adiabatically increasing the intensity of the laser beams to a variable final value V_0 . The calibration of V_0 had an uncertainty of about 20%. A simple linear ramp with a constant rate dV_0/dt of 0.5 $E_F \text{ ms}^{-1}$ was used unless otherwise specified. This satisfies the interband adiabaticity condition of $dV_0/dt \ll 16E_F^2/\hbar$.

Ballistic expansion for the detection of the different momentum components was provided by a magnetic field sequence (shown in Fig. 1) that quickly brought the system out of the strongly interacting regime when all confinement was switched off. During the magnetic field ramp of about 150 μ s, the lattice potential was kept on. The first 2 ms of expansion took place at 470 G, at which the molecules are tightly bound, before the field was ramped back up to 730 G in the next 4.5 ms, at which the weakly bound molecules strongly absorb light near the atomic resonance line and could be observed by absorption imaging. The specific magnetic field sequence was chosen to minimize collisions within technical capabilities.

Received 29 June; accepted 25 August 2006.

- Hofstetter, W., Cirac, J. I., Zoller, P., Demler, E. & Lukin, M. D. High-temperature superfluidity of fermionic atoms in optical lattices. *Phys. Rev. Lett.* **89**, 220407 (2002).
- Scalapino, D. J. The case for $d_{x^2-y^2}$ pairing in the cuprate superconductors. *Phys. Rep.* **250**, 329–365 (1995).
- Greiner, M., Regal, C. A. & Jin, D. S. Emergence of a molecular Bose–Einstein condensate from a Fermi gas. *Nature* **426**, 537–540 (2003).
- Jochim, S. *et al.* Bose–Einstein condensation of molecules. *Science* **302**, 2101–2103 (2003).
- Zwierlein, M. W. *et al.* Observation of Bose–Einstein condensation of molecules. *Phys. Rev. Lett.* **91**, 250401 (2003).
- Regal, C. A., Greiner, M. & Jin, D. S. Observation of resonance condensation of fermionic atom pairs. *Phys. Rev. Lett.* **92**, 040403 (2004).
- Zwierlein, M. W. *et al.* Condensation of pairs of fermionic atoms near a Feshbach resonance. *Phys. Rev. Lett.* **92**, 120403 (2004).
- Zwierlein, M. W., Abo-Shaeer, J. R., Schrirotzek, A., Schunck, C. H. & Ketterle, W. Vortices and superfluidity in a strongly interacting Fermi gas. *Nature* **435**, 1047–1051 (2005).

9. Modugno, G., Ferlaino, F., Heidemann, R., Roati, G. & Inguscio, M. Production of a Fermi gas of atoms in an optical lattice. *Phys. Rev. A* **68**, 011601(R) (2003).
10. Köhl, M., Moritz, H., Stöferle, T., Gunther, K. & Esslinger, T. Fermionic atoms in a three-dimensional optical lattice: observing Fermi surfaces, dynamics, and interactions. *Phys. Rev. Lett.* **94**, 080403 (2005).
11. Stöferle, T., Moritz, H., Gunther, K., Köhl, M. & Esslinger, T. Molecules of fermionic atoms in an optical lattice. *Phys. Rev. Lett.* **96**, 030401 (2006).
12. Volz, T. *et al.* A Mott state of molecules. *Condensed Matt.* (submitted); preprint at <<http://arxiv.org/abs/cond-mat/0605184>> (2006).
13. Winkler, K. *et al.* Repulsively bound atom pairs in an optical lattice. *Nature* **441**, 853–856 (2006).
14. Duan, L.-M. Effective Hamiltonian for Fermions in an optical lattice across a Feshbach resonance. *Phys. Rev. Lett.* **95**, 243202 (2005).
15. Andrews, M. R. *et al.* Observation of interference between two Bose condensates. *Science* **275**, 637–641 (1997).
16. Madison, K. W., Chevy, F., Wohlleben, W. & Dalibard, J. Vortex formation in a stirred Bose–Einstein condensate. *Phys. Rev. Lett.* **84**, 806–809 (2000).
17. Abo-Shaeer, J. R., Raman, C., Vogels, J. M. & Ketterle, W. Observation of vortex lattices in Bose–Einstein condensates. *Science* **292**, 476–479 (2001).
18. Anderson, B. P. & Kasevich, M. A. Macroscopic quantum interference from atomic tunnel arrays. *Science* **282**, 1686–1689 (1998).
19. Greiner, M., Mandel, O., Esslinger, T., Hänsch, T. W. & Bloch, I. Quantum phase transition from a superfluid to a Mott insulator in a gas of ultracold atoms. *Nature* **415**, 39–44 (2002).
20. Schunck, C., Zwierlein, M. W., Schirotzek, A. & Ketterle, W. Superfluid expansion of a strongly interacting Fermi gas. *Condensed Matt.* (submitted); preprint at <<http://arxiv.org/abs/cond-mat/0607298>> (2006).
21. Onofrio, R. *et al.* Observation of superfluid flow in a Bose–Einstein condensed gas. *Phys. Rev. Lett.* **85**, 2228–2231 (2000).
22. Schori, C., Stöferle, T., Moritz, H., Köhl, M. & Esslinger, T. Excitations of a superfluid in a three-dimensional optical lattice. *Phys. Rev. Lett.* **93**, 240402 (2004).
23. Xu, K. *et al.* Sodium Bose–Einstein condensates in an optical lattice. *Phys. Rev. A* **72**, 043604 (2005).
24. Moritz, H., Stöferle, T., Gunter, K., Köhl, M. & Esslinger, T. Confinement induced molecules in a 1d Fermi gas. *Phys. Rev. Lett.* **94**, 210401 (2005).
25. Koetsier, A. O., Dickerscheid, D. B. M. & Stoof, H. T. C. BEC–BCS crossover in an optical lattice. *Condensed Matt.* (submitted); preprint at <<http://arxiv.org/abs/cond-mat/0604186>> (2006).
26. Gubbels, K. B., Dickerscheid, D. B. M. & Stoof, H. T. C. Dressed molecules in an optical lattice. *Condensed Matt.* (submitted); preprint at <<http://arxiv.org/abs/cond-mat/0605056>> (2006).
27. Hadzibabic, Z. *et al.* Fifty-fold improvement in the number of quantum degenerate fermionic atoms. *Phys. Rev. Lett.* **91**, 160401 (2003).
28. Stan, C. A. & Ketterle, W. Multiple species atom source for laser-cooling experiments. *Rev. Sci. Instrum.* **76**, 063113 (2005).
29. Bartenstein, M. *et al.* Precise determination of ⁶Li cold collision parameters by radio-frequency spectroscopy on weakly bound molecules. *Phys. Rev. Lett.* **94**, 103201 (2004).
30. Gould, P. L., Ruff, G. A. & Pritchard, D. E. Diffraction of atoms by light: the near-resonant Kapitza–Dirac effect. *Phys. Rev. Lett.* **56**, 827–830 (1986).

Acknowledgements We thank E. Demler, Z. Hadzibabic and M. Zwierlein for discussions. This work was supported by the NSF, the Office of Naval Research and NASA.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.K.C. (jitkee@mit.edu).

Exploration of molecular dynamics during transient sorption of fluids in mesoporous materials

Rustem Valiullin¹, Sergej Naumov¹, Petrik Galvosas¹, Jörg Kärger¹, Hyung-June Woo², Fabien Porcheron³ & Peter A. Monson³

In recent years, considerable progress has been made in the development of novel porous materials with controlled architectures and pore sizes in the mesoporous range^{1–4}. An important feature of these materials is the phenomenon of adsorption hysteresis: for certain ranges of applied pressure, the amount of a molecular species adsorbed by the mesoporous host is higher on desorption than on adsorption, indicating a failure of the system to equilibrate. Although this phenomenon has been known for over a century, the underlying internal dynamics responsible for the hysteresis remain poorly understood^{5–9}. Here we present a combined experimental and theoretical study in which microscopic and macroscopic aspects of the relaxation dynamics associated with hysteresis are quantified by direct measurement and computer simulations of molecular models. Using nuclear magnetic resonance techniques^{10–14} and Vycor porous glass^{15,16} as a model mesoporous system, we have explored the relationship between molecular self-diffusion and global uptake dynamics. For states outside the hysteresis region, the relaxation process is found to be essentially diffusive in character; within the hysteresis region, the dynamics slow down dramatically and, at long times, are dominated by activated rearrangement of the adsorbate density within the host material.

The measurement of gas adsorption is a classical characterization technique³ that remains of enormous importance in the development of new kinds of porous materials with applications in areas ranging from separation and catalysis to microelectronics⁴. The shape of an adsorption isotherm reflects the interactions of the adsorbed molecules with the porous material surfaces, as well as the effect of confinement on the adsorbed fluid properties. For mesoporous materials, an important feature of such measurements is the occurrence of hysteresis at temperatures below the bulk critical temperature. This is widely assumed to be associated with the formation of a condensed liquid-like state^{3,5}, and represents a possible example of the influence of spatial confinement on a phase transition^{5,17}. However, recent mean-field density functional theory studies of a disordered lattice-gas model suggest that hysteresis can occur even though the system does not exhibit an underlying phase transition^{6,8}. The results of these studies suggest a central role for the internal dynamics of such systems in giving rise to hysteresis.

An important feature of the dynamics of these systems is suggested by the free energy landscape in the hysteresis region that was predicted in theoretical studies^{6,8}. For a medium density state of an adsorbed fluid in a material such as Vycor porous glass, there are a very large number of spatial arrangements of the fluid that are consistent with the average density. These states of the confined fluid represent local minima of the free energy, and they are metastable. The approach to equilibrium (a global minimum in the free energy)

at longer times involves transitions between these states via activated barrier crossings, as shown by dynamic Monte Carlo simulations for the lattice model of a fluid in Vycor porous glass⁸. This is an intrinsically slower process than the relaxation through diffusive mass transfer that characterizes the behaviour away from the hysteresis region. The hysteresis is similar to that observed in disordered ferromagnets as modelled by the random-field Ising model¹⁸, and the dynamical behaviour resembles that predicted for this model^{19,20}. In contrast to the abundance of experimental data on adsorption equilibria and hysteresis^{1,3,5}, there is a striking lack of experiments probing the dynamics. In this Letter, for the first time to our knowledge, a self-consistent set of experimental data on dynamics and transport is presented and used to test the conclusions emerging from these theoretical studies.

We have used nuclear magnetic resonance (NMR) as a probe of the adsorbed fluid^{10,11}, which allows us to study both the dynamics of adsorption and desorption and the self-diffusivity. Self-diffusion characterizes random translational motion of molecules under equilibrium conditions. It is driven by the internal kinetic energy of the molecules, making it sensitive to various molecular interactions. The self-diffusion coefficient D relates molecular mean-squared displacements $\langle r^2 \rangle$ to the observation time t via Einstein's relation $\langle r^2 \rangle = 6Dt$. In porous materials, D is affected by the restrictions imposed by confinement²¹. Importantly, an inhomogeneous distribution of the adsorbate inside the porous material may also affect its dynamical properties. The spatial variation of the adsorbed fluid density leads to varying local diffusivities. The experimentally measured diffusivity is a quantity averaged over distances exceeding by far the typical size of the inhomogeneities, which is appropriate for the present work. Thus, it may be exploited as a parameter characterizing the adsorbate state in mesopores under different conditions^{13,14}. Transient sorption, on the other hand, reflects collective dynamics of molecules under non-equilibrium conditions²¹.

Among various experimental techniques for studying self-diffusion, pulsed field gradient NMR is the most suitable for fluids in porous materials^{11,21}. This method is based on the creation of an initial coherence of the nuclei and on following its loss in an applied magnetic field gradient due to displacements of the nuclei. Thus, the characteristics of molecular motion may be directly followed over a broad timescale from milliseconds to seconds. At the same time, the initial coherence leads to an observable NMR free induction decay (FID) signal, which is proportional to the number of nuclei in the sample, that is, the amount adsorbed. Thus, transient uptake and molecular diffusion may be probed simultaneously.

In Fig. 1 are shown the adsorption isotherm for cyclohexane in Vycor porous glass and the corresponding diffusivities measured at different points of the isotherm. The amount adsorbed is given in the

¹Abteilung Grenzflächenphysik, Fakultät für Physik und Geowissenschaften, Universität Leipzig, D-04103 Leipzig, Germany. ²Department of Chemistry, University of Nevada, Reno, Nevada 89557, USA. ³Department of Chemical Engineering, University of Massachusetts, Amherst, Massachusetts 01003, USA.

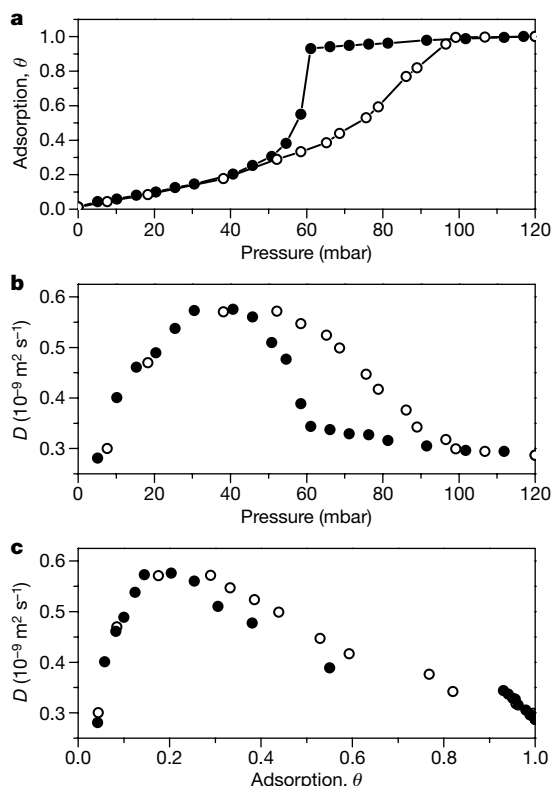


Figure 1 | Experimental adsorption isotherm and effective diffusivities of cyclohexane in Vycor porous glass. **a, b,** Normalized adsorption θ (see Methods for details) of cyclohexane in pores of Vycor porous glass (**a**) and the diffusivity D (**b**) measured as a function of external vapour pressure P on the adsorption (open symbols) and the desorption (filled symbols) branches of the isotherm at $T = 297$ K. Error bars (s.d.; not shown) are of the order of two symbol heights. The diffusivities were measured using an observation time $t_d = 10$ ms. No dependence of D on t_d in the range 2–300 ms was observed. **c,** Combination of the data of **a** and **b** to show D as a function of concentration.

fraction θ of complete pore filling (details are given in the Methods section) and is referred to as concentration. The observed hysteresis (Fig. 1a) is of the so-called H2 type³, which is typical of Vycor porous glass. As the diffusivities depend on concentrations, and

concentrations depend on pressure, a plot of diffusivities versus pressure should likewise exhibit hysteretic behaviour (Fig. 1b). Following refs 13 and 14, the observed diffusivities may be satisfactorily predicted by a simple gas kinetics approach involving both surface diffusion and Knudsen diffusion in the free pore space. Remarkably, on plotting D versus θ (Fig. 1c), we see that the hysteretic behaviour is preserved. Evidently, depending on the history through which a particular state was attained—even at equal pore concentrations in the region of the irreversible sorption—the dynamical properties of the adsorbate can be different. In this sense, different values of D at other equal conditions (temperature, concentration) may be considered as a manifestation of the history-dependent adsorbate distribution.

In addition to these experiments, we have made calculations of the adsorption and self-diffusion for the lattice gas model of a fluid in Vycor (see Methods for details), following the procedure given in refs 8 and 22. The resulting isotherms and diffusivities for two different surface–fluid interaction strengths are shown in Fig. 2. Notably, the experimentally observed trends in Fig. 1 are qualitatively well reproduced in the simulations (a more quantitative agreement might be achieved by adjusting interaction parameters in the model, but it is the qualitative behaviour that concerns us here). The effect of the surface interaction on the resulting patterns of the diffusivity is discernible. Apparently, the experiments and model calculations yield a qualitatively consistent picture of the equilibrium dynamics, which we can use as a basis for analysing the transient behaviour.

In Fig. 3, the results of the transient sorption measurements are correlated with the information from the diffusion studies. The figure presents typical examples of the uptake kinetics following a relatively small (5 mbar) stepwise change of the external vapour pressure (Fig. 3a refers to the out-of-hysteresis region and Fig. 3b to the hysteresis region). With the independently known diffusivities and by assuming that diffusion is the rate-limiting process, one may find the uptake function, $\theta_{\text{diff}}(t)$, as a solution of Fick's law. For cylindrical geometry of the monolithic Vycor particle used in the present study, the solution is given by²³

$$\theta_{\text{diff}}(t) = \theta_0 + (\theta_{\text{eq}} - \theta_0) \left(1 - \frac{4}{a^2} \sum_{n=1}^{\infty} \frac{1}{\alpha_n^2} \exp\{-\alpha_n^2 D t\} \right) \quad (1)$$

where a represents the radius of the cylindrical particle, α_n are the positive roots of the equation $J_0(\alpha a) = 0$, θ_0 and θ_{eq} are the initial and the final concentrations, and J_0 is the Bessel function of the first kind. With θ_{eq} taken from the experimental data in Fig. 3 at long times

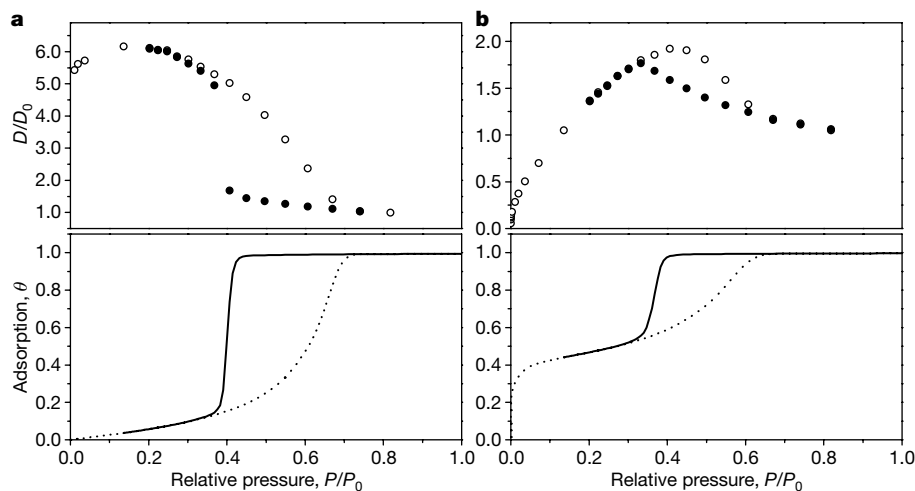


Figure 2 | Simulated adsorption/desorption isotherm and effective diffusivities for a lattice model of a fluid confined in a Vycor porous glass. Lower graphs, adsorption/desorption isotherms (adsorption, dotted line; desorption, solid line), and upper graphs, self-diffusion coefficients (adsorption, open circles; desorption, filled circles), for the lattice model of a

fluid in a Vycor model, calculated by Monte Carlo simulations. **a,** Data for a weak surface interaction ($\gamma = 1$); **b,** data for a strong surface interaction ($\gamma = 3$). Here γ is the ratio of the strengths of the solid–fluid and fluid–fluid nearest neighbour interactions in the lattice model.

where the uptake is almost complete, and with D from Fig. 1, $\theta_{\text{diff}}(t)$ is readily calculated, as shown in Fig. 3 by the dotted lines. Importantly, equation (1) with θ_0 and θ_{eq} chosen as described correctly reproduces the experimental data for the region out of hysteresis, but fails in the hysteresis region. Slower equilibration in the hysteresis region has been noted before^{9,24,25} but has been interpreted in terms of a decrease in the diffusivity^{24,25}, based on comparing experimental uptake curves with simple micro-kinetic models. However, what we are seeing here is a fundamental difference in the nature of the relaxation dynamics for states within the hysteresis region compared with those out of this region, and this is not simply a change in the rate of diffusion.

The mechanism leading to the anomalous behaviour in the hysteresis region can be illuminated using the lattice gas model of a fluid confined in Vycor porous glass⁸. Figure 4 shows results from dynamic Monte Carlo simulations of the adsorption versus time for a system quenched into the hysteresis regime, and is in very good agreement with the experimental observation. Detailed analysis of the accompanying data obtained from these calculations allowed us to identify a two-stage mechanism of the transient uptake in the hysteresis region, which may be visualized in the following way. Adsorption at low pressures consists of diffusion of the gas into the pore structure, with the formation of adsorbed layers on the pore walls. In general, the entire pore space is accessible to mass transfer from the bulk, making the dynamics purely diffusive. As the pressure

increases, the average density builds up until it becomes possible for condensation processes to occur, in which liquid 'droplets' appear at various places in the pore structure, forming bridges between the pore walls. The formation of these 'droplets' is initially not driven by a global equilibration process so that the distribution of the fluid density around the system represents a local minimum of the free energy.

Subsequent density redistribution and/or droplet growth are required for relaxation to global equilibrium (at a given external gas pressure), and these are essentially activated processes requiring the crossing of barriers of height¹⁹ $b\xi^\psi > kT$ (where ξ is the characteristic 'droplet' size, $\psi > 1$, k is Boltzmann's constant, T is temperature, and b is a constant determined by the fluid properties and external conditions). A quantitative description of the relaxation dynamics is suggested by analysis of the critical dynamics of the random-field Ising system^{19,20}. The lattice gas model used here can be interpreted as a kind of random-field Ising model with site-dilution⁸. The total relaxation can be represented by a linear combination of two relaxation processes: a diffusive one given by equation (1) with the relative weight K and an activated one²⁰ with the weight $(1 - K)$, so that

$$\theta(t) = K\theta_{\text{diff}}(t) + (1 - K)(\theta_{\text{eq}} - (\theta_{\text{eq}} - \theta_0) \exp\{-[\ln(t/\tau_0)/\ln(\tau_a/\tau_0)]^p\}) \quad (2)$$

where τ_0 is the characteristic microscopic time, τ_a is the typical relaxation time for activated dynamics, and p is the parameter reflecting distribution of the barrier heights^{26,27}. Such a two-term function with $p = 3$ (refs 26, 27) can be used to give an excellent fit of the $\theta(t)$ results from both experiment and our Monte Carlo simulations. Similar relaxation processes have also been observed in studies of liquid-liquid phase separation in Vycor porous glass²⁷.

In conclusion, the extremely slow relaxation inherent to activated dynamics governs the approach to equilibrium in the hysteresis region at longer times. This explains why hysteresis, although representing a departure from equilibrium, is so reproducible in experiments. The convergence of experiment and theory shown here represents a major step forward in understanding the fundamentals of relaxation phenomena in confined fluids. At the same time, correlating diffusion and relaxation in porous materials represents an important approach for the exploration of the internal dynamics in mesoscopic systems. Continuation of these studies with tailor-made materials, spanning the total range of pore architecture between the

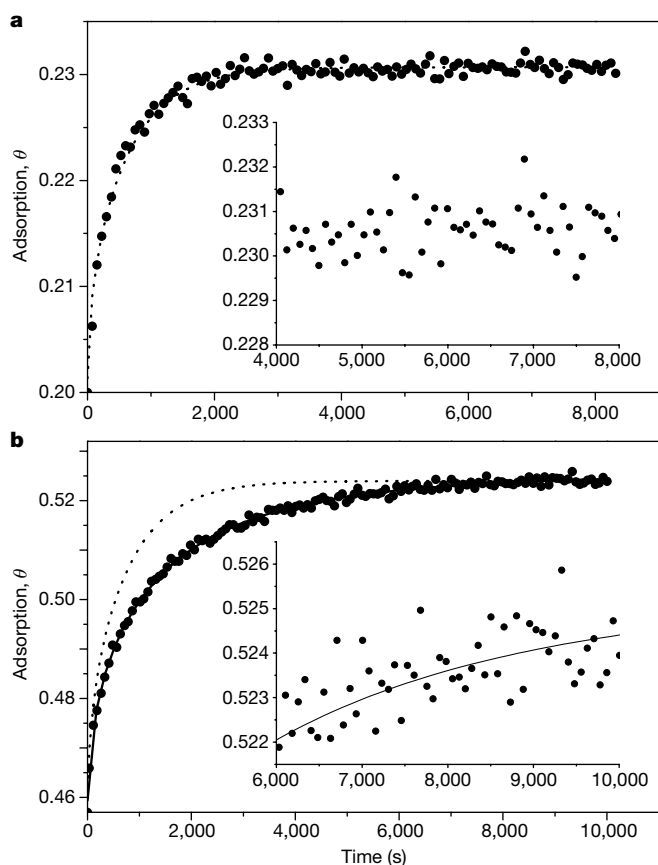


Figure 3 | Experimental transient sorption of cyclohexane in Vycor porous glass. **a**, **b**, Typical adsorption kinetic data (points) obtained upon stepwise change of the external vapour pressure from 40 mbar to 45 mbar (**a**) and 70 mbar to 75 mbar (**b**). Insets show the long-time part of the same data; axis quantities and units are the same as the main figures. The experimental data of the adsorption branch of the isotherm shown in Fig. 1 represent the θ_{eq} values at the longest observation time ($t \approx 10^4$ s). The dotted lines in the figures show the kinetics calculated using equation (1). The solid line in **b** and in the inset shows results from equation (2) with parameters $\tau_0 = 600$ s, $\tau_a = 4,540$ s, $K = 0.8$ and $p = 3$.

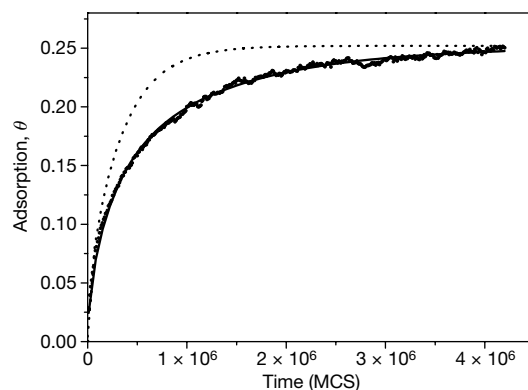


Figure 4 | Simulated transient sorption for a lattice model of a fluid confined in Vycor porous glass. Graph shows density versus time from a dynamic Monte Carlo simulation of the uptake of a fluid into a Vycor glass in the hysteresis regime. Simulations have been performed for a weak surface interaction $\gamma = 0.7$, where γ is the ratio of the strengths of the solid-fluid and fluid-fluid nearest-neighbour interactions in the lattice model. Time is given in Monte Carlo steps (MCS). The broken line is a fit to the short-time behaviour (diffusion model). The solid line is from equation (2) with parameters $\tau_0 = 1.4 \times 10^5$ MCS, $\tau_a = 2.6 \times 10^6$ MCS, $K = 0.68$ and $p = 3$.

limiting cases of straight channels¹⁴ and hollow-sphere arrangements²⁸, should lead to a range of additional insights, including a more complete understanding of the effect of the porous material geometry on the relaxation dynamics.

METHODS

NMR experiments. The experimental set-up used in the present study consists of an NMR glass tube containing the porous material, connected through a valve to a large reservoir with the bulk vapour phase of the fluid under study. As a porous host, a Vycor porous glass particle in the form of a cylinder of 3 mm diameter and of 12 mm length has been used. The internal porous structure of the particle consists of a disordered network of interconnected pores with an effective diameter of about 6 nm (refs 15, 16). The volume of the gas reservoir significantly exceeds that of the porous material. The vapour pressure can be varied from millibars to the saturated vapour pressure P_0 of the liquid under study. After setting a certain pressure in the reservoir, opening the valve causes a stepwise change of the external conditions for the porous material. A subsequent equilibration of the intra-pore density was followed by recording the NMR FID signal intensity S_{FID} as a function of time. In order to avoid the necessity of calibrating the measured NMR signal, we confine ourselves to presenting a normalized value, the relative intra-pore concentration $\theta = [S_{\text{FID}}(P)]/[S_{\text{FID}}(P_0)]$, where $S_{\text{FID}}(P_0)$ is the signal intensity measured at full saturation of the pores. The self-diffusion coefficients D were measured at sufficiently long times after the pressure change when no notable change of the amount adsorbed was observed.

Lattice model simulations. In the model used here, the positions of the adsorbed molecules and the porous solid are given by occupying the sites of a cubic lattice (either simple cubic or face-centred cubic). The positions of the solid sites are fixed, and chosen by coarse-graining the solid distribution obtained from the gaussian random field model of the interfacial structure generated in the phase separation process occurring in the preparation of the porous glass^{8,29}. The adsorption/desorption isotherm was calculated using Monte Carlo simulations³⁰ in the grand canonical ensemble. The self-diffusion was simulated using Kawasaki dynamics, a process in which molecules hop to adjacent sites on the lattice following the Metropolis criterion for accepting the new configurations generated³⁰.

Received 29 March; accepted 14 August 2006.

- Schüth, F., Sing, K. S. W. & Weitkamp, J. (eds). *Handbook of Porous Solids* (Wiley-VCH, Weinheim, 2002).
- Kresge, C. T., Leonowicz, M. E., Roth, W. J., Vartuli, J. C. & Beck, J. S. Ordered mesoporous molecular sieves synthesized by a liquid-crystal template mechanism. *Nature* **359**, 710–712 (1992).
- Sing, K. S. W., Rouquerol, F. & Rouquerol, J. *Adsorption by Powders and Solids* (Academic, London, 1999).
- Barton, T. J. *et al.* Tailored porous materials. *Chem. Mater.* **11**, 2633–2656 (1999).
- Gelb, L. D., Gubbins, K. E., Radhakrishnan, R. & Sliwinski-Bartkowiak, M. Phase separation in confined systems. *Rep. Prog. Phys.* **62**, 1573–1659 (1999).
- Kierlik, E., Monson, P. A., Rosinberg, M. L., Sarkisov, L. & Tarjus, G. Capillary condensation in disordered porous materials: Hysteresis versus equilibrium behavior. *Phys. Rev. Lett.* **87**, 055701 (2001).
- Bocquet, L., Charlaix, E., Ciliberto, S. & Crassous, J. Moisture-induced ageing in granular media and the kinetics of capillary condensation. *Nature* **396**, 735–737 (1998).
- Woo, H. J. & Monson, P. A. Phase behavior and dynamics of fluids in mesoporous glasses. *Phys. Rev. E* **67**, 041207 (2003).
- Wallacher, D., Kunzner, N., Kovalev, D., Knorr, N. & Knorr, K. Capillary condensation in linear mesopores of different shape. *Phys. Rev. Lett.* **92**, 195704 (2004).

- Kärger, J., Pfeifer, H. & Heink, W. Principles and application of self-diffusion measurements by nuclear magnetic resonance. *Adv. Magn. Reson.* **12**, 2–89 (1988).
- Callaghan, P. T. *Principles of Nuclear Magnetic Resonance Microscopy* (Clarendon, Oxford, 1991).
- Beyea, S. D., Caprihan, A., Glass, S. J. & DiGiovanni, A. Nondestructive characterization of nanopore microstructure: Spatially resolved Brunauer-Emmett-Teller isotherms using nuclear magnetic resonance imaging. *J. Appl. Phys.* **94**, 935–941 (2003).
- Valiullin, R., Kortunov, P., Kärger, J. & Timoshenko, V. Concentration-dependent self-diffusion of liquids in nanopores: A nuclear magnetic resonance study. *J. Chem. Phys.* **120**, 11804–11814 (2004).
- Valiullin, R., Kortunov, P., Kärger, J. & Timoshenko, V. Concentration-dependent self-diffusion of adsorbates in mesoporous materials. *Magn. Reson. Imaging* **23**, 209–214 (2005).
- Levitz, P., Ehret, G., Sinha, S. K. & Drake, J. M. Porous Vycor glass: The microstructure as probed by electron microscopy, direct energy transfer, small-angle scattering, and molecular adsorption. *J. Chem. Phys.* **95**, 6151–6161 (1991).
- Kikkinides, E. S. *et al.* Combination of small angle scattering and three-dimensional stochastic reconstruction for the study of adsorption-desorption processes in Vycor porous glass. *J. Chem. Phys.* **112**, 9881–9887 (2000).
- Evans, R. Fluids adsorbed in narrow pores — phase-equilibria and structure. *J. Phys. Condens. Matter* **2**, 8989–9007 (1990).
- Sethna, J. P. *et al.* Hysteresis and hierarchies: dynamics of disorder-driven first-order phase transformations. *Phys. Rev. Lett.* **70**, 3347–3350 (1993).
- Fisher, D. S. Scaling and critical slowing down in random-field Ising systems. *Phys. Rev. Lett.* **56**, 416–419 (1986).
- Huse, D. A. Critical dynamics of random-field Ising systems with conserved order parameter. *Phys. Rev. B* **36**, 5383–5387 (1987).
- Kärger, J. & Ruthven, D. M. *Diffusion in Zeolites and Other Microporous Solids* (Wiley & Sons, New York, 1992).
- Woo, H. J., Sarkisov, L. & Monson, P. A. Mean-field theory of fluid adsorption in a porous glass. *Langmuir* **17**, 7472–7475 (2001).
- Carslaw, H. S. & Jaeger, J. C. *Conduction of Heat in Solids* (Clarendon, Oxford, 1946).
- Rajniak, P., Soos, M. & Yang, R. T. Unified network model for adsorption-desorption in systems with hysteresis. *Am. Inst. Chem. Eng. J.* **45**, 735–750 (1999).
- Lee, J. W., Shim, W. G. & Moon, H. Adsorption equilibrium and kinetics for capillary condensation of trichloroethylene on MCM-41 and MCM-48. *Micropor. Mesopor. Mater.* **73**, 109–119 (2004).
- Ogielski, A. T. & Huse, D. A. Critical behavior of the three-dimensional dilute Ising antiferromagnet in a field. *Phys. Rev. Lett.* **56**, 1298–1301 (1986).
- Dierker, S. B. & Wiltzius, P. Random-field transition of a binary liquid in a porous-medium. *Phys. Rev. Lett.* **58**, 1865–1868 (1987).
- Smarsly, B. *et al.* Microstructural characterization of polystyrene-block-poly(ethylene oxide)-templated silica films with cubic-ordered spherical mesopores. *Langmuir* **19**, 7295–7301 (2003).
- Levitz, P. Off-lattice reconstruction of porous media: critical evaluation, geometrical confinement and molecular transport. *Adv. Colloid Interf. Sci.* **77**, 71–106 (1998).
- Newman, M. E. J. & Barkema, G. T. *Monte Carlo Methods in Statistical Physics* (Oxford Univ. Press, Oxford, 1999).

Acknowledgements Research on this project at the University of Massachusetts was supported by the National Science Foundation. R.V. and J.K. thank the German Science Foundation and the Alexander von Humboldt Foundation for support.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to R.V. (valiullin@uni-leipzig.de).

A palaeotemperature curve for the Precambrian oceans based on silicon isotopes in cherts

François Robert¹ & Marc Chaussidon²

The terrestrial sediment record indicates that the Earth's climate varied drastically in the Precambrian era (before 550 million years ago), ranging from surface temperatures similar to or higher than today's to global glaciation events¹. The most continuous record of sea surface temperatures of that time has been derived from variations in oxygen isotope ratios of cherts (siliceous sediments)², but the long-term cooling of the oceans inferred from those data^{3–5} has been questioned because the oxygen isotope signature could have been reset through the exchange with hydrothermal fluids after deposition of the sediments⁶. Here we show that the silicon isotopic composition of cherts more than 550 million years old shows systematic variations with age that support the earlier conclusion of long-term ocean cooling and exclude post-depositional exchange as the main source of the isotopic variations. In agreement with other lines of evidence^{1,7}, a model of the silicon cycle in the Precambrian era shows that the observed silicon isotope variations imply seawater temperature changes from about 70 °C 3,500 million years ago to about 20 °C 800 million years ago.

We selected a suite of chert samples that covers most of the available sedimentary record from the early Archaean era (3.5 Gyr ago) to the present (see Methods). Our sample set includes 9 Phanerozoic samples and 99 Precambrian samples dating from about 3.5 Gyr ago to about 0.5 Gyr ago and coming from 23 geological formations (Supplementary Table S1). These cherts were selected on the basis of several criteria ($\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ values⁸) that were indicative of good preservation of their original structure and composition (some of them have been extensively studied for their abundant microfossils⁹). The O and Si isotopic composition were measured with a multicollector Cameca IMS 1270 ion microprobe (see Methods).

The $\delta^{18}\text{O}$ values of cherts show large variations that, for a given geological age, range from the maximum previously determined by Knauth and Lowe³ (hereafter reported as $\delta^{18}\text{O}_{\text{KL}}$) down to $\delta^{18}\text{O} = \delta^{18}\text{O}_{\text{KL}} - 15\text{‰}$ (Fig. 1). The curve defined by the maximum $\delta^{18}\text{O}$ ($\delta^{18}\text{O}_{\text{KL}}$) for a given age was proposed to reflect formation temperatures at depth in the sediment. These formation temperatures are in turn related to past seawater temperatures (Fig. 1). The oxygen isotopic fractionation between chert and water ($\Delta^{18}\text{O} = \delta^{18}\text{O}_{\text{chert}} - \delta^{18}\text{O}_{\text{water}}$) depends on temperature according to the equation $1,000\ln(\Delta^{18}\text{O}) = (3.09 \times 10^6 T^{-2}) - 3.29$. The large range in $\delta^{18}\text{O}$ indicates that a significant fraction of Precambrian cherts had their original $\delta^{18}\text{O}$ values lowered by post-depositional isotopic exchange with meteoric or hydrothermal fluids³ or, in some cases, were formed over a large range of temperature⁵.

The $\delta^{30}\text{Si}$ values of the present Phanerozoic cherts (nine samples) vary from $-1.7 \pm 0.8\text{‰}$ to $+1.3 \pm 0.3\text{‰}$ (1 s.d.) (see Supplementary Table S1), covering the range previously determined for diatoms collected from open ocean surface waters (from $+0.9\text{‰}$ to $+1.9\text{‰}$)¹⁰, diatoms from sedimentary cores ($+1.3\text{‰}$ to $+1.6\text{‰}$)¹¹

or siliceous sediments from black smokers (-3.5‰ to -0.2‰)^{12,13}. In contrast, the present set of Precambrian cherts shows much higher $\delta^{30}\text{Si}$ values (ranging from $-1.1 \pm 0.4\text{‰}$ to $+5.0 \pm 0.8\text{‰}$; Supplementary Table S1). Such high $\delta^{30}\text{Si}$ values have no counterpart in the Phanerozoic sedimentary record but were previously reported for a few Precambrian stromatolitic cherts^{12,13}. Detailed studies of Onverwacht cherts⁴ or of modern deep-sea cherts¹⁴ have shown that their $\delta^{18}\text{O}$ values can depart from equilibrium with sea water by up to about -10‰ ⁵ because of either the contribution of meteoric waters to the diagenetic fluids or the range in the crystallization temperatures during burial diagenesis. To account in our data set for this oxygen isotopic variability inherent in the process of chert formation, we tentatively consider here that cherts having $\delta^{18}\text{O} \leq \delta^{18}\text{O}_{\text{KL}} - 6\text{‰}$ did not preserve an original isotopic oceanic signature. Using this criterion, it turns out that all the cherts (56 out of 99) having $\delta^{18}\text{O}_{\text{KL}} - 6\text{‰} \leq \delta^{18}\text{O} \leq \delta^{18}\text{O}_{\text{KL}}$ show both an increase in their $\delta^{30}\text{Si}$ values with time from 3.5 to 0.8 Gyr (Fig. 2a) when they return abruptly to modern values and a corresponding positive correlation between their $\delta^{30}\text{Si}$ and $\delta^{18}\text{O}$ values (Fig. 2b). Note that using a

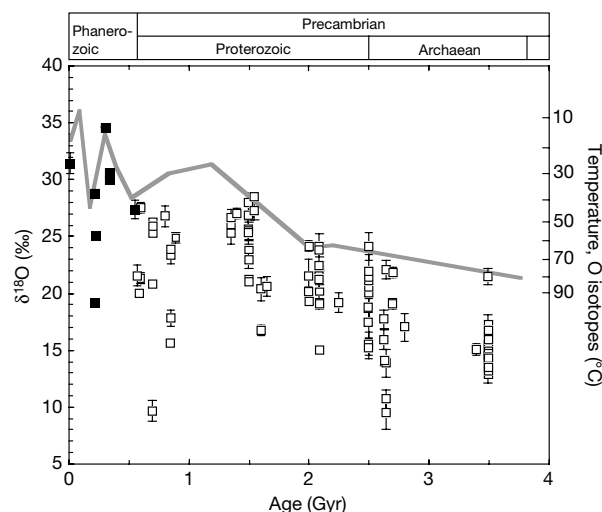


Figure 1 | Variation in chert $\delta^{18}\text{O}$ values with geological age. Open squares, our Precambrian samples; filled squares, Phanerozoic samples. Error bars correspond to 1σ error on the mean of three to five analyses per sample. The solid line (referred to in the text as $\delta^{18}\text{O}_{\text{KL}}$ (ref. 3)), shows the highest $\delta^{18}\text{O}$ value found in cherts at a given age. Considering that the ice caps were absent—that is, $\delta^{18}\text{O}_{\text{sea water}} = -1\text{‰}$ —variation in $\delta^{18}\text{O}_{\text{KL}}$ with time may reflect seawater temperature variations³ between about 70 °C in the early Archaean and about 30 °C in the late Proterozoic. $\delta^{18}\text{O}$ values of modern deep-sea cherts¹⁴ can depart from equilibrium with sea water by up to about -4‰ (see the text). These variations, which are inherent in the process of chert crystallization, control the precision on seawater temperature.

¹Muséum National d'Histoire Naturelle, CNRS LEME NanoAnalyses, UMS 2679, 57 rue Cuvier, 75005 Paris, France. ²Centre de Recherches Pétrographiques et Géochimiques, CRPG-CNRS BP20, 54501 Vandœuvre-lès-Nancy, France.

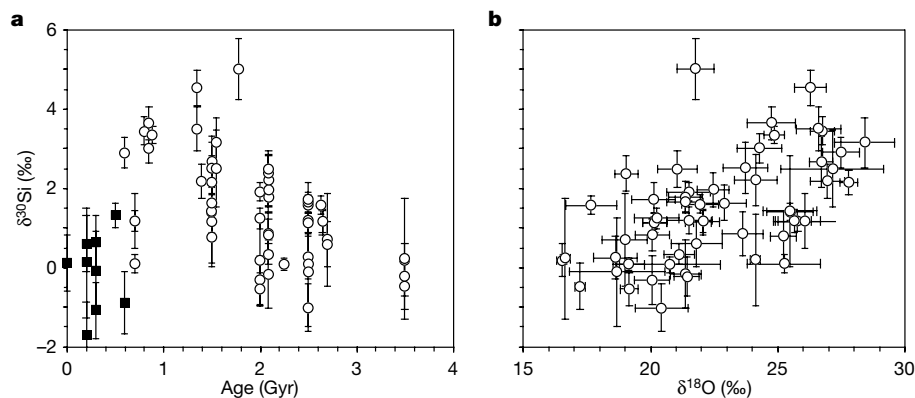


Figure 2 | Variation of $\delta^{30}\text{Si}$ with $\delta^{18}\text{O}$ values and geological age for samples with the highest $\delta^{18}\text{O}$ values. The highest $\delta^{18}\text{O}$ values for a given age are defined by $\delta^{18}\text{O}_{\text{KL}} - 6\text{‰} < \delta^{18}\text{O} < \delta^{18}\text{O}_{\text{KL}}$, see text. Error bars for $\delta^{18}\text{O}$ and $\delta^{30}\text{Si}$ values correspond to 1σ error on the mean of three to five analyses per sample. Precambrian cherts (open circles) have $\delta^{30}\text{Si}$ values

threshold of 4‰ ($\delta^{18}\text{O}_{\text{KL}} - 4\text{‰} \leq \delta^{18}\text{O} \leq \delta^{18}\text{O}_{\text{KL}}$) restricts the number of chert samples to 40 out of 99 but does not change the correlations observed in Fig. 2. The positive correlation between $\delta^{30}\text{Si}$ and $\delta^{18}\text{O}$ vanishes for cherts with $\delta^{18}\text{O} < \delta^{18}\text{O}_{\text{KL}} - 6\text{‰}$ (Fig. 3). Thus, the threshold at about 6‰ on $\delta^{18}\text{O}$ values seems a relevant criterion to distinguish between samples with different O–Si isotopic systematics.

The $\delta^{30}\text{Si}$ – $\delta^{18}\text{O}$ correlation (Fig. 2b) seems hardly compatible with post-depositional isotopic exchange with hydrothermal fluids, which would have lowered both $\delta^{18}\text{O}$ and $\delta^{30}\text{Si}$. If a low temperature (about 15°C) is assumed for the Precambrian oceans, cherts must form with $\delta^{18}\text{O} = +32\text{‰}$ and, to fit the correlation reported in Fig. 2, with a $\delta^{30}\text{Si}$ of between $+6\text{‰}$ and $+8\text{‰}$. In addition, samples with $\delta^{18}\text{O} < \delta^{18}\text{O}_{\text{KL}} - 6\text{‰}$, for instance Rietgat samples (Supplementary Table S1), show that the $\delta^{30}\text{Si}$ values remain essentially unchanged whereas $\delta^{18}\text{O}$ is reset through the interaction with high-temperature fluids (Fig. 3). This behaviour is qualitatively in agreement with the fact that, as Si is insoluble in fluids, the water/rock ratios for the two isotopic systems are quite different.

In contrast with oxygen isotopes, there is apparently no significant temperature effect on the Si isotopic fractionation between precipi-

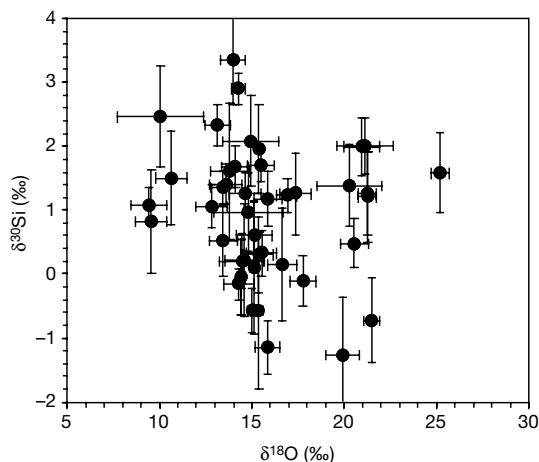


Figure 3 | Variations of $\delta^{30}\text{Si}$ with $\delta^{18}\text{O}$ values for samples with low $\delta^{18}\text{O}$ values ($\delta^{18}\text{O} < \delta^{18}\text{O}_{\text{KL}} - 6\text{‰}$). Error bars for $\delta^{18}\text{O}$ and $\delta^{30}\text{Si}$ values correspond to 1σ error on the mean of three to five analyses per sample. The lack of correlation between $\delta^{18}\text{O}$ and $\delta^{30}\text{Si}$ values for these samples shows that the decrease in $\delta^{18}\text{O}$ values (by up to 15‰), probably caused by the interaction with high-temperature fluids, has a limited effect on the $\delta^{30}\text{Si}$ values.

much higher than those of Phanerozoic ones (filled squares). The change in $\delta^{30}\text{Si}$ values is sharp at the end of the Precambrian. In addition, the $\delta^{30}\text{Si}$ values increase during the Precambrian (a), which parallels the $\delta^{18}\text{O}$ evolution, resulting in a positive correlation between $\delta^{30}\text{Si}$ and $\delta^{18}\text{O}$ values (b).

tated and dissolved silicon^{12,15,16}. Thus, variations of chert $\delta^{30}\text{Si}$ values must primarily reflect variations in the $\delta^{30}\text{Si}$ of the fluids from which they formed.

The major difference in the marine silica cycle between the Phanerozoic¹⁷ and the Precambrian¹⁸ is the rapid development about 0.6 Gyr ago of organisms that used dissolved silicon to build their skeleton: biogenic silica precipitation maintains the Si concentration in sea water today at $[\text{Si}]_{\text{sw}} \approx 4.5 \text{ p.p.m.}$ ¹⁷. In contrast, in the Precambrian $[\text{Si}]_{\text{sw}}$ was probably controlled by both the solubility of amorphous silica, which resulted in a much higher $[\text{Si}]_{\text{sw}}$ than today (for example 49 p.p.m. at 20°C (ref. 21)), and the rate of silicification of the oceanic crust¹⁸. This major difference has profound consequences for the $\delta^{30}\text{Si}$ value of Precambrian oceans and for its evolution with time.

In the absence of biogenic precipitation of silica—that is, during the Precambrian—it can be considered that seawater Si had two major outputs: the sedimentary silica precipitated from pore sea water and the hydrothermal silica resulting from the silicification of the oceanic crust. The sedimentary silica was transformed into cherts during diagenesis, process that we consider to be at the origin of our chert samples with $\delta^{18}\text{O} \leq \delta^{18}\text{O}_{\text{KL}} - 6\text{‰}$. At steady state, isotopic mass balance must be fulfilled between the input and output Si fluxes. This Si budget can be written in a manner similar to that formulated, for instance, for the oceanic carbon isotopic budget, which is controlled by the relative outputs of carbonates and reduced carbon and the isotopic fractionation between the two carbon pools. For Si, the budget should be controlled by the relative outputs of sedimentary and hydrothermal silica and the associated isotopic fractionations (see Supplementary Fig. S2) so that

$$\delta^{30}\text{Si}_{\text{In}} = f\delta^{30}\text{Si}_{\text{Sed}} + (1 - f)\delta^{30}\text{Si}_{\text{Hvd}}$$

where $\delta^{30}\text{Si}_{\text{In}}$ is the isotopic composition of the Si input (input includes contributions from mantle and continental crust) to the marine cycle, f is the relative fraction of sedimentary Si precipitated from sea water or pore sea water in sediments (subscript Sed) and $(1 - f)$ is the counterpart retrieved by the hydrothermal silicification of the crust (subscript Hyd). The major Si isotopic fractionation in this system (defined as $\Delta^{30}\text{Si}_{\text{Hvd-In}} = \delta^{30}\text{Si}_{\text{Hvd}} - \delta^{30}\text{Si}_{\text{In}}$) is predicted to have taken place during the silicification of the crust. Although not yet studied experimentally, a negative value seems likely for $\Delta^{30}\text{Si}_{\text{Hvd-In}}$: first, $\delta^{30}\text{Si}$ values down to -3.5‰ (ref. 13) have been reported for siliceous sediments precipitated at black smokers from hydrothermal fluids with $\delta^{30}\text{Si} = -0.3\text{‰}$ (ref. 11) and second, isotopic fractionations between precipitated silica and dissolved silicon are observed to be of the order of -1‰ (refs 12, 15, 16). In addition,

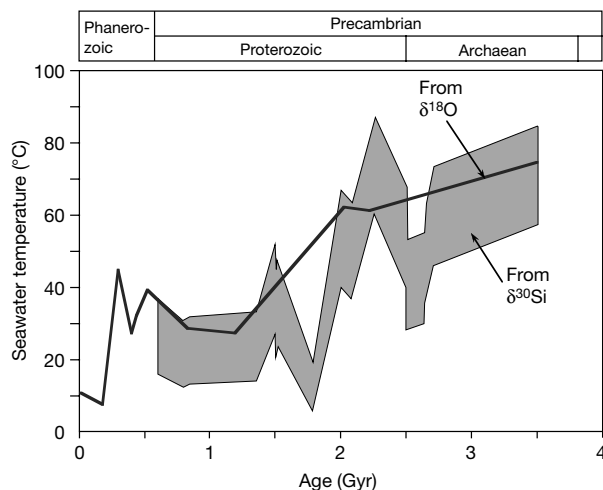


Figure 4 | Variations in oceanic temperatures modelled from the $\delta^{30}\text{Si}$ values from cherts (grey area) compared with the curve proposed from $\delta^{18}\text{O}$ values³. The silicon isotope thermometer has been calculated on the assumption that either the maximum temperature (at 3.5 Gyr ago) or the minimum temperature (at 0.8 Gyr ago) indicated by the $\delta^{18}\text{O}$ values is correct (see Supplementary Information). Although the global significance of the silicon isotope temperature curve can be questioned at certain geological times (namely, apparent strong cooling at about 1.8 and about 2.5 Gyr ago) as a result of limited sampling, the $\delta^{30}\text{Si}$ values from the cherts are consistent with a significant decrease in oceanic temperatures since 3.5 Gyr ago, in agreement with results from oxygen isotopes.

if the silicification process involves the diffusion of Si, then kinetic isotopic effects can be predicted that would enrich the hydrothermally deposited silica in ^{28}Si . Because the exact value of $\Delta^{30}\text{Si}_{\text{Hyd-In}}$ is unknown, we considered it to be a free parameter calculated by the model. Numerical calculations (detailed in the Supplementary Information) show that it varies between about -1‰ and about -3‰ ; that is, within a range that agrees with theoretical expectations. It is obvious from the mass balance equation that, for negative $\Delta^{30}\text{Si}_{\text{Hyd-In}}$ values, $\delta^{30}\text{Si}_{\text{Sed}}$ is positive and increases when f decreases. This is in qualitative agreement with the positive $\delta^{30}\text{Si}$ found in the present cherts.

Furthermore it seems that the regular increase in chert $\delta^{30}\text{Si}$ values during the Precambrian (Fig. 2a) can be used as a 'palaeothermometer' for sea water: this comes from the fact that seawater temperature (T_{sw}) is the major control on f .

Considering that the Si concentration in the hydrothermal fluids circulating in the silicified crust is at equilibrium with quartz at the fluid temperature (T_{Hyd}), whereas sea water is in equilibrium with amorphous silica at T_{sw} . When hydrothermal fluids percolating the oceanic crust return to the surface, they have to cool and to equilibrate with sea water (or with pore sea water). Because the solubility of quartz in hydrothermal fluids¹⁹ at temperatures for which the precipitation rate of silica is maximal²⁰, $T > 190^\circ\text{C}$ (see Supplementary Information), is higher than that of amorphous silica²¹ for any realistic seawater temperature (that is, $T < 70^\circ\text{C}$), these fluids lose a large fraction of their Si on the way to the surface. This fraction, referred to as $(1 - f)$ in the mass balance equation, depends on the difference between T_{Hyd} and T_{sw} , which implies that ultimately it depends on T_{sw} : if T_{sw} decreases, the amount of Si sequestered by silicification of the crust increases (that is, f decreases) and thus $\delta^{30}\text{Si}_{\text{Sed}}$ increases.

This process holds only if part of the silicon injected into seawater precipitates as amorphous silica. This would occur systematically if dissolved silicon in sea water were at, or slightly above, the saturation level relative to amorphous silica (see Supplementary Information) as postulated for the Proterozoic¹⁸. After the appearance of Si-fixing organisms, presumably in the early Phanerozoic, it can be supposed that, as today, sea water was always maintained strongly undersaturated relative to amorphous silica. In this situation, hydrothermal

fluids that are discharged into sea water cool and are diluted, always remaining undersaturated relative to amorphous silica. In our model, this would correspond to $f = 0$, in agreement with the rather constant and low (close to mantle) $\delta^{30}\text{Si}$ values found for Phanerozoic cherts (Fig. 2a).

As detailed numerically in Supplementary Information, this approach allows us to calculate—from the $\delta^{30}\text{Si}$ -age relationship defined by Precambrian cherts—a relationship between T_{sw} and age. For this calculation, the $\delta^{30}\text{Si}$ values of the chert samples with $\delta^{18}\text{O} \leq \delta^{18}\text{O}_{\text{KL}} - 6\text{‰}$ have been averaged for each geological formation (Supplementary Table S2). The $\delta^{30}\text{Si}$ of the input flux was taken to be -0.3‰ ; this is the value of the mantle and of the fluids at black smokers^{11,22}; that is, the value of the mantle input to the sea-water-crust system. Continental clay formation and silicification processes^{23,24} that could have changed the $\delta^{30}\text{Si}$ value of the riverine input are ignored in this model. They could not account for the $\delta^{30}\text{Si}$ - $\delta^{18}\text{O}$ correlation (Fig. 2b) because the $\delta^{18}\text{O}$ value of the oceans in the Precambrian was not controlled by continental erosion.

Taking 70 and 30°C , given by $\delta^{18}\text{O}_{\text{KL}}$, for seawater temperatures at 3.5 and 0.85 Gyr, respectively, allows us to solve the mass balance equation for a range of $\Delta^{30}\text{Si}_{\text{Hyd-In}}$ and T_{Hyd} (see Supplementary Information for details). The corresponding range of possible T_{sw} values for the Precambrian is indicated by the grey region in Fig. 4. The calculated trend of decreasing T_{sw} during the Precambrian is then in striking agreement with that predicted previously from O isotopes. This approach also predicts the positive correlation observed in cherts between $\delta^{18}\text{O}$ and $\delta^{30}\text{Si}$. Thus, O and Si isotopic compositions of cherts concur to show that large changes in surface temperature occurred during the Precambrian.

METHODS

Geological formations (location, geology and radiometric ages) of all the samples studied here are reviewed in refs 9, 25. Specific information on aliquots of the samples studied can be found in refs 25–28. The other samples come from three different collections: first, the Caltech collection (courtesy of P. Knauth and S. Epstein; referred to in Supplementary Table S1 as K for Knauth samples; previous isotopic studies of some of the samples studied can be found in refs 2, 3, 29); second, the samples collected by S. Awramik (courtesy of S. Awramik; referred to as 'n of month-day-year' of the recovery); and third, the Precambrian Palaeobiology Research Group collection referred to as PPRG (courtesy of J. W. Schopf)²⁵.

The sample fragments were embedded in epoxy (about ten in each mount), coated with gold and sputtered with a Cs^+ primary beam (5–10 nA intensity and about $30\text{ }\mu\text{m}$ in size) and the electron gun. The negative $^{16}\text{O}^-$, $^{18}\text{O}^-$, $^{28}\text{Si}^-$ and $^{30}\text{Si}^-$ ions produced during sputtering were accelerated at 10 kV and analysed at a mass resolution $M/\Delta M$ of about 3,000 in multi-collection mode by using Faraday cups (O and Si isotopic ratios were measured in different sessions; see recent developments of these techniques^{23,30}). We used quartz standards (Rose quartz and NBS-28) and an internal chert standard (miocene chert from the Paris basin) that was systematically included in all the mounts and analysed together with the samples. For Si, we assumed that there was no matrix effect on instrumental mass fractionation between chert and quartz.

$\delta^{30}\text{Si}$ and $\delta^{18}\text{O}$ are expressed relative to the NBS-28 and Vienna standard mean ocean water (VSMOW) international standards, respectively; for example $\delta^{30}\text{Si} = 1,000 \times [({}^{30}\text{Si}/{}^{28}\text{Si})/({}^{30}\text{Si}/{}^{28}\text{Si})_{\text{NBS}} - 1]$ and similarly for $\delta^{18}\text{O}$.

During the different analytical sessions in which the two quartz standards were run together, the Rose quartz standard gave $\delta^{30}\text{Si}$ values relative to NBS-28 of $+0.24 \pm 0.14\text{‰}$, $+0.20 \pm 0.20\text{‰}$, $+0.16 \pm 0.40\text{‰}$ and $+0.34 \pm 0.40\text{‰}$. The precision of the $\delta^{18}\text{O}$ and $\delta^{30}\text{Si}$ measurements was not limited by counting statistics (about $\pm 0.1\text{‰}$ and $\pm 0.05\text{‰}$, respectively) but probably by sample heterogeneity: our internal chert standard (Miocene chert) gave $\delta^{18}\text{O}$ and $\delta^{30}\text{Si}$ values ranging within $\pm 0.49\text{‰}$ (1 σ , 53 analyses) and $\pm 0.28\text{‰}$ (1 σ , 48 analyses), respectively. Between three and five analytical spots were made systematically on each sample to document the possible isotopic variability on a small scale (for example, analysis of silica veins cross-cutting the silica ground mass were avoided). When available, the bulk $\delta^{18}\text{O}$ values of some samples—previously determined at the milligram scale by the classical BrF_5 oxidation procedure^{2,3,29}—were compared with the values obtained by the ion probe at the nanogram scale: both sets of data agreed within $\pm 1.9\text{‰}$ on average (Supplementary Fig. S1). This rather large scatter illustrates the fact that, in some rare cases, oxygen isotopic heterogeneity up to $+3.3\text{‰}$ (sample K500 in

Supplementary Table S1) are present at the millimetre scale. The data (reported in Supplementary Table S1) are averages over the different analytical spots made on a given sample. The 1σ error (from $\pm 0.2\text{‰}$ to $\pm 2.3\text{‰}$ for $\delta^{18}\text{O}$ and from $\pm 0.25\text{‰}$ to $\pm 1.4\text{‰}$ for $\delta^{30}\text{Si}$) on these averages reflects the isotopic variability within a given sample.

Received 8 December 2005; accepted 8 September 2006.

- Lowe, D. R. & Tice, M. N. Geologic evidences for Archean atmospheric and climatic evolution: fluctuating levels of CO_2 , CH_4 , and O_2 with an overriding tectonic control. *Geology* **32**, 493–496 (2004).
- Knauth, L. P. & Epstein, S. Hydrogen and oxygen isotope ratios in nodular and bedded cherts. *Geochim. Cosmochim. Acta* **40**, 1095–1108 (1976).
- Knauth, L. P. & Lowe, D. R. Oxygen isotope geochemistry of cherts from the Onverwacht Group (3.4 billion years) Transvaal, South Africa, with implications for secular variations in the isotopic composition of cherts. *Earth Planet. Sci. Lett.* **41**, 209–222 (1978).
- Knauth, L. P. & Lowe, D. R. High Archean climatic temperature inferred from oxygen isotope geochemistry of cherts in the 3.5 Ga Swaziland Supergroup, South Africa. *Geol. Soc. Am. Bull.* **115**, 566–580 (2003).
- Knauth, L. P. Temperature and salinity history of the Precambrian ocean: implications for the course of microbial evolution. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **219**, 53–69 (2005).
- Degens, E. T. & Epstein, S. Relationship between $^{18}\text{O}/^{16}\text{O}$ ratios in coexisting carbonates, cherts and diatomites. *Bull. Am. Assoc. Petrol. Geol.* **46**, 534–542 (1962).
- Kasting, J. F. Methane and climate during the Precambrian era. *Precamb. Res.* **137**, 119–129 (2005).
- Beaumont, V. & Robert, F. Nitrogen isotope ratios of kerogens in Precambrian cherts: A record for the evolution of atmosphere chemistry? *Precamb. Res.* **96**, 63–82 (1999).
- Schopf, J. W. & Klein, C. (eds). *The Proterozoic Biosphere* (Cambridge Univ. Press, New York, 1992).
- De La Rocha, C. L., Brzezinski, M. A. & DeNiro, M. Silicon-isotope composition of diatoms as an indicator of past oceanic change. *Nature* **395**, 680–683 (1998).
- De La Rocha, C. L., Brzezinski, M. A. & DeNiro, M. A first look at the distribution of the stable isotopes of silicon in natural waters. *Geochim. Cosmochim. Acta* **61**, 2467–2477 (2000).
- Ding, T. P. *et al.* *Silicon Isotope Geochemistry* (Geological Publishing House, Beijing, 1996).
- Jiang, S.-Y., Palmer, M. R., Peng, Q.-M. & Yang, J.-H. Chemical and stable isotopic compositions of Proterozoic metamorphosed evaporites and associated tourmalines from Houxianyu borate deposit, eastern Liaoning, China. *Chem. Geol.* **135**, 189–211 (1997).
- Kolodny, Y. & Epstein, S. Stable isotope geochemistry of deep sea cherts. *Geochim. Cosmochim. Acta* **40**, 1195–1209 (1976).
- De La Rocha, C. L., Brzezinski, M. A. & DeNiro, M. Fractionation of silicon isotopes by marine diatoms during biogenic silica formation. *Geochim. Cosmochim. Acta* **61**, 5051–5056 (1997).
- Ding, Y., Wan, D., Wang, C. & Zhang, F. F. Silicon isotope compositions of dissolved silicon suspended in the Yangtze River, China. *Geochim. Cosmochim. Acta* **68**, 205–216 (2004).
- Tréguer, P. *et al.* The silica balance in the world ocean: a reestimate. *Science* **268**, 375–379 (1995).
- Siever, R. The silica cycle in the Precambrian. *Geochim. Cosmochim. Acta* **56**, 3265–3272 (1992).
- Von Damm, K. L., Bischoff, J. L. & Rosenbauer, R. J. Quartz solubility in hydrothermal seawater—an experimental study and equation describing quartz solubility for up to 0.5N NaCl solutions. *Am. J. Sci.* **291**, 977–1007 (1991).
- Rimstidt, J. D. & Barnes, H. L. The kinetics of silica water reactions. *Geochim. Cosmochim. Acta* **44**, 1683–1699 (1980).
- Gunnarsson, I. & Arnorsson, S. Amorphous silica solubility and the thermodynamic properties of H_4SiO_4 in the range of 0° to 350°C at P_{sat} . *Geochim. Cosmochim. Acta* **64**, 2295–2307 (2000).
- Douthitt, C. B. The geochemistry of the stable isotopes of silicon. *Geochim. Cosmochim. Acta* **46**, 1449–1458 (1982).
- Basile-Doelsch, I., Meunier, J.-D. & Parron, C. Another continental pool in the terrestrial silicon cycle. *Nature* **433**, 399–402 (2005).
- Ziegler, K., Chadwick, O. A., Brzezinski, M. A. & Kelly, E. F. Natural variations of $\delta^{30}\text{Si}$ ratios during progressive basalt weathering, Hawaiian Islands. *Geochim. Cosmochim. Acta* **69**, 4597–4610 (2005).
- Schopf, J. W. (ed.). *Earth's Earliest Biosphere: its Origin and Evolution* (Princeton Univ. Press, Princeton, New Jersey, 1983).
- Allison, C. W. & Awramik, S. Organic-walled microfossils from earliest Cambrian or latest Proterozoic Tindir Group rocks, Northwest Canada. *Precamb. Res.* **43**, 253–294 (1989).
- Karpeta, W. P. Bedded cherts in the Rietgat Formation, Hartbeesfontein, South Africa: a late Archean to early Proterozoic magadiitic alkaline playa lake deposit? *South Afr. J. Geol.* **92**, 29–36 (1989).
- Awramik, S. M. *et al.* Prokaryotic and eukaryotic microfossils from a Proterozoic–Phanerozoic transition in China. *Nature* **315**, 655–658 (1985).
- Karhu, J. & Epstein, S. The implication of the oxygen isotope records in coexisting cherts and phosphates. *Geochim. Cosmochim. Acta* **50**, 1745–1756 (1986).
- Rollion-Bard, C., Chaussidon, M., France-Lanord, Ch. pH control on oxygen isotopic composition of symbiotic corals. *Earth Planet. Sci. Lett.* **215**, 275–288 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank P. Knauth for constructive and critical reviews, which greatly helped in clarifying the text with regard to the qualitative significance of the thermometric equations proposed for silicon isotopes. We thank W. Schopf, S. Awramik, P. Knauth and the late S. Epstein for providing an unlimited and disinterested access to the most representative samples of their own collections.

Author Contributions Both authors contributed equally to this work.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to F.R. (robert@mnhn.fr) or M.C. (chocho@crpg.cnrs-nancy.fr).

Hydrous olivine unable to account for conductivity anomaly at the top of the asthenosphere

Takashi Yoshino¹, Takuya Matsuzaki¹, Shigeru Yamashita¹ & Tomoo Katsura¹

The oceanic asthenosphere is observed to have high electrical conductivity, which is highly anisotropic in some locations^{1,2}. In the directions parallel and normal to the plate motion, the conductivity is of the order of 10^{-1} and 10^{-2} S m⁻¹, respectively, which cannot be explained by the conductivity of anhydrous olivine². But because hydrogen can be incorporated in olivine at mantle pressures^{3–5}, this observation has been attributed to olivine hydration, which might cause anisotropically high conductivity by proton migration^{1,2,6,7}. To examine this hypothesis, here we report the effect of water on electrical conductivity and its anisotropy for hydrogen-doped and undoped olivine at 500–1,500 K and 3 GPa. The hydrous olivine has much higher conductivity and lower activation energy than anhydrous olivine in the investigated temperature range. Nevertheless, extrapolation of the experimental results suggests that conductivity of hydrous olivine at the top of the asthenosphere should be nearly isotropic and only of the order of 10^{-2} S m⁻¹. Our data indicate that the hydration of olivine cannot account for the geophysical observations², which instead may be explained by the presence of partial melt elongated in the direction of plate motion.

The electrical conductivity of hydrogen-bearing nominally anhydrous minerals has not been previously determined owing to the difficulty of reliable measurement; hydrogen inside silicate minerals can easily escape under high temperature conditions (above 1,000 K) because of its high diffusivity. Recently, electrical conductivity of hydrous wadsleyite and ringwoodite has been reported, based on measurements over quite short times at a single temperature condition for each run in order to minimize water loss⁸. However, this method cannot determine a reliable activation energy because of the absence of systematic measurements over a wide temperature range. There are two approaches to overcoming the above problem: one is to measure conductivity in a closed system for H₂O, and the other is to measure conductivity at low temperatures at which the hydrogen diffusion rate is sufficiently low. The former is impractical, because of the difficulty of sealing-in H₂O by using an insulating material. Although electrical conductivity measurements under low temperature conditions (<1,000 K) require high insulation resistance, we have measured the electrical conductivity of hydrous olivine at temperatures of 500–1,000 K using low frequency (0.01 Hz) a.c. signals. To clarify the effect of hydrogen on the electrical conductivity, we have conducted conductivity measurements both for initially hydrogen-doped samples and for undoped samples.

The conductivity measurements were conducted at 3 GPa in a Kawai-type multi-anvil apparatus along several heating–cooling cycles in each run (see Supplementary Methods). Figure 1 is an Arrhenius plot showing conductivities of the hydrogen-doped and the undoped olivines along the three principal axes. The hydrogen-doped olivine has a higher conductivity (by two orders of magnitude) than the undoped olivine. The conductivity of the hydrogen-doped

olivine is anisotropic in the investigated temperature range; it is higher along the [100] crystallographic axis than the [010] and [001] axes when data are normalized to the same hydrogen content, based on a linear relationship between the hydrogen concentration and the conductivity. The trend of anisotropy is consistent with that of hydrogen diffusion in olivine⁷. The conductivity of the undoped olivine can be divided into two regions. At lower temperatures, the conductivity has a smaller temperature dependence than it has at higher temperatures. In each temperature region, the logarithmic conductivity increases linearly with decreasing reciprocal temperature. The conductivity in the high temperature region is consistent with that of anhydrous olivine previously obtained⁹.

Electrical conductivities σ can be expressed as follows:

$$\sigma = \sigma_0 \exp\left(-\frac{E}{kT}\right) \quad (1)$$

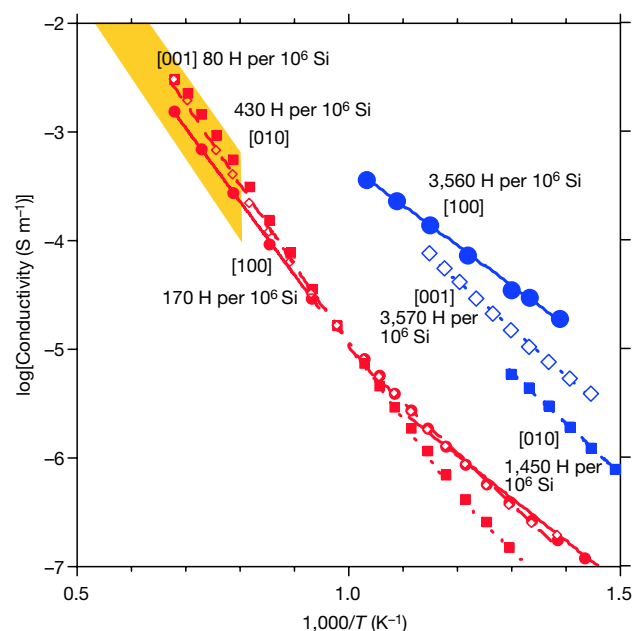


Figure 1 | Electrical conductivity of olivine as a function of reciprocal temperature. The symbols indicate raw data for each sample with different hydrogen contents. Red and blue lines indicate results for the undoped and the hydrogen-doped samples, respectively. Filled circles, filled squares and open diamonds represent the different crystallographic directions: [100], [010] and [001], respectively. Solid, dotted and dashed lines indicate the data fitting for [100], [010] and [001] directions, respectively. The yellow area indicates the range of conductivity of anhydrous olivine at 4 GPa reported previously⁹.

¹Institute for Study of the Earth's Interior, Okayama University, Misasa, Tottori 682-0193, Japan.

Table 1 | Summary of experimental runs

Run no.	Axis	C _{OH} [*]	C _{OH} [†]	σ_{oh} [‡]	E_h [‡] (eV)	σ_{op} [§]	E_p [§] (eV)	Remarks
1K520	[100]	0(20)	168(58)	65.04	1.35	0.055	0.78	
1K524	[010]	0(12)	433(156)	169.9	1.42	0.384	0.98	
1K521	[001]	0(7)	82(47)	286.2	1.47	0.250	0.89	
1K471	[100]	3,560(187)	3,503(189)			2.329	0.73	
1K482	[010]	1,450(291)	700(224)			7.416	0.93	Dehydration
1K478	[001]	3,569(128)	1,118(124)			7.666	0.87	Dehydration

All experiments were conducted at 3 GPa.

^{*} Hydrogen concentration (in H per 10⁶ Si) of synthesized hydrous olivine before the conductivity measurement experiment, determined by FTIR spectroscopy. Parentheses denote one standard deviation for five analyses.

[†] Hydrogen concentration (in H per 10⁶ Si) after the conductivity measurement experiment.

[‡] Pre-exponential factor (σ_{oh}) and activation energy (E_h) for hopping conductivity in the higher temperature range.

[§] Pre-exponential factor (σ_{op}) and activation energy (E_p) for proton conductivity of hydrous olivine. For the relatively dry olivine, they were calculated in the lower temperature range.

where σ_0 is a pre-exponential factor, E is activation energy, k is the Boltzmann constant and T is temperature. The pre-exponential term and activation energy are summarized in Table 1. The activation energies of the undoped olivine in the high temperature region are around 1.4 eV, which is consistent with the previous studies^{9–11}. In this temperature region, transport of small polarons should be the dominant conduction mechanism. The activation energies of the hydrogen-doped olivine along each crystallographic axis are found to be $E_{[100]} = 0.73$, $E_{[010]} = 0.97$ and $E_{[001]} = 0.87$ eV. These values are essentially identical to those of the undoped olivine in the lower temperature region: $E_{[100]} = 0.78$, $E_{[010]} = 0.98$ and $E_{[001]} = 0.89$ eV. These observations suggest that the conduction mechanism of the undoped olivine in the lower temperature region should be attributed to migration of hydrogen (proton conduction). In fact, the undoped olivine contained a small amount (80–400 H per 10⁶ Si) of hydrogen during the electrical conductivity measurements (Table 1). We suggest that the intrinsic (small polaron) conductivity of olivine is masked by the extrinsic (proton) conductivity, owing to the difference of the activation energy between these two mechanisms. The fact that the activation energy of the proton conduction for the [100] direction is lower than for the other axes suggests that the anisotropy of the proton conductivity becomes small, and [010] rather than [100] becomes the highest conductive direction, at high temperatures.

The activation energies for proton conduction are generally lower than those obtained from proton diffusion experiments⁷. At the temperature range of 1,073–1,273 K, the proton conductivity is nearly isotropic, whereas the proton diffusion in the [100] direction is significantly faster than that in the other directions⁷. These discrepancies may imply that the experimental series of diffusion and conduction measurements are controlled by different migration mechanisms of protons in olivine crystal. The hydrogen concentrations for the conduction experiments were nearly in equilibrium, whereas those for the diffusion experiments were not in equilibrium (for details, see Supplementary Discussion). On the other hand, hydrogen diffusion data for wadsleyite¹² also show that the activation energy for proton diffusion (0.92 eV) is clearly higher than that for proton conduction (0.66 eV)⁸. This tendency is same as for the case of olivine. Therefore, mechanisms of proton migration for these two phenomena could be essentially different. In other words, proton diffusion measurements are not necessarily an appropriate way to estimate proton conductivity.

To assess the presence of water in the upper mantle, we compare the present experimental data with the electromagnetic study underneath the Pacific Ocean; we do this by calculating the logarithmic conductivity of hydrous olivine versus depth, based on the following assumptions (Fig. 2). A potential temperature of 1,623 K was assumed for the adiabatic geotherm. Activation energies for proton conductivity were assumed to be constant at temperatures higher than the present measurement range. The upper limit of the electrical conductivity for hydrous olivine down to a depth of 410 km was calculated on the basis of the reported pressure dependence of the maximum H₂O solubility in olivine¹³. A linear relation between

conductivity and amount of hydrogen was assumed, based on the Nernst–Einstein relation¹⁴. The oxygen fugacity in the mantle was assumed to be the Ni/NiO buffer¹⁵, although that in the sub-ridge mantle is two orders of magnitude less than Ni/NiO (ref. 16).

The electrical conductivity in the oceanic upper mantle above 400 km depth varies from $\sim 10^{-3}$ to 10^{-1} S m⁻¹ (refs 1, 2, 17, 18). The one-dimensional electrical conductivity structure in the mantle of the one-fourth of the Earth beneath the north Pacific Ocean obtained by a semi-global electromagnetic induction study¹⁷ is quite consistent with the laboratory-based conductivity–depth profile predicted from the dry olivine⁹. In contrast, the deep electrical conductivity profile, obtained from analysis of data from a submarine cable

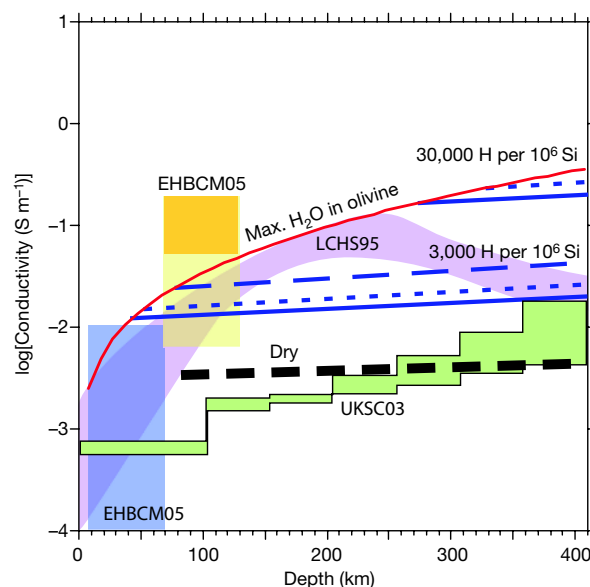


Figure 2 | Comparison of laboratory data on electrical conductivity of hydrous olivine as a function of H₂O content with the geophysically observed electrical conductivity in the upper mantle beneath the Pacific.

The red line indicates the upper bound of the electrical conductivity of the hydrous olivine. Because the maximum hydroxyl solubility as a function of depth based on equation (1) of ref. 12 is three times the hydrogen content based on a method of ref. 5, we applied one-third the value of the hydrogen content used by ref. 12 as the hydrogen content of olivine. The blue solid line, long-dashed line and short-dashed line represent the electrical conductivity of hydrous olivine in the [100], [010] and [001] directions, respectively. The electrical conductivity of hydrous olivine for each crystallographic direction is normalized to 3,000 and 30,000 H per 10⁶ Si, respectively. XSD00⁹ (thick black dashed line) shows the electrical conductivity of dry olivine. The green and pink areas (UKSC03¹⁷ and LCHS95¹⁸) represent the geophysically observed conductivity profile in the Pacific. The orange and light-orange regions (EHBCM05²) indicate the range of conductivity in directions parallel to and normal to the plate motion around a depth of 100 km near the East Pacific Rise. The light blue area (EHBCM05²) represents the range of conductivity above 60 km near the East Pacific Rise.

extending from Hawaii to North America, showed a conductivity peak of 10^{-1} S m^{-1} at the depth range 200–250 km (ref. 18). This profile is quite consistent with the upper limit of the conductivity predicted from the maximum H_2O solubility in olivine. Therefore, the high-conductivity anomaly in a part of the mantle underneath the Pacific Ocean can be interpreted as the presence of $\sim 10^4 \text{ H per } 10^6 \text{ Si}$, which is less than the storage capacity of H_2O in olivine. The variation in the conductivity–depth profiles may be derived from the local heterogeneity of the H_2O distribution in the mantle.

A recent magnetotelluric study at the southern East Pacific Rise reported two important features of the electrical conductivity profile near a mid-ocean ridge². The conductivity largely increases with increasing depth from 50 to 100 km. It also becomes strongly anisotropic in this depth range. Around 100 km, the difference of conductivities between the direction parallel (10^{-1} S m^{-1}) and normal (10^{-2} S m^{-1}) to the plate motion reaches one order of magnitude. This remarkable change was interpreted as an increase in H_2O content in olivine in this depth range. However, the present results cannot support the hydrous olivine hypothesis. First, the estimated conductivity of hydrous olivine at a depth of 100 km can be only as high as 10^{-2} S m^{-1} in this region. Although this value is distinctly higher than that of dry olivine^{9–11}, it is not enough to explain the maximum conductivity ($\sim 0.1 \text{ S m}^{-1}$) obtained from the East Pacific Rise. To explain the conductivity of 0.1 S m^{-1} , a hydrogen concentration of over 17,000 H per 10^6 Si , which exceeds the maximum solubility of H_2O in olivine at 3 GPa (ref. 5), is required. Second, extrapolation of our data to high temperatures shows that anisotropy of hydrous olivine becomes small (less than by a factor of 2) at the temperature condition corresponding to the top of the asthenosphere. The weak anisotropy disagrees with the one-order-of-magnitude difference of the conductivity between ridge-parallel and ridge-normal directions.

In addition, seismic observations showed high shear velocity in the direction parallel to plate motion. The seismic anisotropy was interpreted as horizontal alignment of the olivine [100] axis, which has the highest shear wave velocity among the crystallographic axes, in the direction of spreading¹⁹. The concentration of [100] to the shear direction is consistent with the olivine fabric deformed under the dry condition. However, the present results suggest that the conductivity of olivine is lowest in the [100] direction. Hence, if the hydrous olivine is oriented in the direction of the [100] axis, as suggested by the seismic study, conductivity in the direction of the plate motion should be lowest, which is opposite to the results of the magnetotelluric study. Increase of the H_2O content may change the olivine fabric²⁰. Taking the crystallographic direction [010] with the highest conductivity at higher temperature conditions into account, it is unlikely that the olivine fabric could explain the highest conductivity direction, even if plastic deformation occurs under both dry ([100] parallel to the shear direction) and wet ([001] parallel to the shear direction) conditions²⁰. These discrepancies lead us to conclude that the geophysically observed high conductivity anomaly² cannot be attributed to the conductivity of hydrous olivine.

An alternative explanation for the high conductivity anomaly is that the conductivity is controlled by a presence of partial melt in this region. Basaltic melt itself has higher electrical conductivity ($>10^{-0.5} \text{ S m}^{-1}$)²¹, and the wetting properties of the basaltic melt against olivine allow the interconnection of partial melt²². Thus, partial melting of peridotites could cause the high conductivity anomaly²³. The top of the asthenosphere ($\sim 60 \text{ km}$) is a region where partial melting of dry peridotite occurs beneath a mid-ocean ridge²⁴. For a normal geotherm with a potential temperature of 1,623 K, partial melting of peridotite is, however, unlikely to occur at the deeper zone ($\sim 100 \text{ km}$) and at locations far from the ridge. Some factors such as water²⁵, carbon dioxide²⁶ and a potassium-rich composition²⁷ should be considered as lowering the melting temperature. For example, the presence of 400 p.p.m. (by weight) of H_2O in peridotite can cause partial melting¹³.

The top of the asthenosphere along the normal geotherm could allow the existence of partial melt, even though the amount of the components lowering the melting temperature is small. If the lithosphere is actually impermeable to the partial melt², partial melt could be trapped at the top of the asthenosphere. It is possible that a small amount of melt is successively added to the top of the asthenosphere from the deeper region, because the flow of mantle convection near the ridge axis should have a considerably large horizontal dimension^{19,28}. The electromagnetic and seismic data obtained from the region just beneath the ridge axis, showing the presence of a narrow melt conduit^{29,30}, may support the presence of an impermeable layer at the top of the asthenosphere.

The high anisotropy of conductivity in the direction of plate motion² can be explained by anisotropic interconnection of melt in the partial melting hypothesis. A recent electromagnetic observation depicted a high-conductivity zone with high anisotropy in the vertical direction beneath the East Pacific Rise²⁹. The origin of this zone is attributed to vertical melt interconnection caused by the material flow. Therefore, we can expect that at the top of the asthenosphere in the region near the ridge, the partial melt should be better connected in the direction of plate motion, which will cause the observed anisotropy of conductivity.

Received 6 December 2005; accepted 22 August 2006.

- Karato, S. The role of hydrogen in the electrical conductivity of the upper mantle. *Nature* **347**, 272–273 (1990).
- Evans, R. L. et al. Geophysical evidence from the MELT area for compositional controls on oceanic plates. *Nature* **437**, 249–252 (2005).
- Bai, Q. & Kohlstedt, D. L. Substantial hydrogen solubility in olivine and implications for water storage in the mantle. *Nature* **357**, 672–674 (1992).
- Bell, D. R. & Rothman, G. R. Water in the earth's mantle: the role of nominally anhydrous minerals. *Science* **255**, 1391–1397 (1992).
- Kohlstedt, D. L., Keppler, H. & Rubie, D. C. Solubility of water in the α , β and γ phases of $(\text{Mg,Fe})_2\text{SiO}_4$. *Contrib. Mineral. Petrol.* **123**, 345–357 (1996).
- Mackwell, S. J. & Kohlstedt, D. L. Diffusion of hydrogen in olivine: implications for water in mantle. *J. Geophys. Res.* **95**, 5079–5088 (1990).
- Kohlstedt, D. L. & Mackwell, S. J. Diffusion of hydrogen and intrinsic point defects in olivine. *Z. Phys. Chem.* **207**, 147–162 (1998).
- Huang, X., Xu, Y. & Karato, S. Water content in the transition zone from electrical conductivity of wadsleyite and ringwoodite. *Nature* **434**, 746–749 (2005).
- Xu, Y., Shankland, T. J. & Duba, A. G. Pressure effect on electrical conductivity of mantle olivine. *Phys. Earth Planet. Inter.* **118**, 149–161 (2000).
- Shankland, T. J. & Duba, A. G. Standard electrical conductivity of isotropic, homogeneous olivine in the temperature range 1200°–1500°C. *Geophys. J. Int.* **103**, 25–31 (1990).
- Wanamaker, B. J. & Duba, A. G. Electrical conductivity of San Carlos olivine along [100] under oxygen- and pyroxene-buffered conditions and implications for defect equilibria. *J. Geophys. Res.* **98**, 489–500 (1993).
- Hae, R., Ohtani, E., Kubo, T., Koyama, T. & Utada, H. Hydrogen diffusivity in wadsleyite and water distribution in the mantle transition zone. *Earth Planet. Sci. Lett.* **243**, 141–148 (2006).
- Hirschmann, M. M., Aubaud, C. & Withers, A. C. Storage capacity of H_2O in nominally anhydrous minerals in the upper mantle. *Earth Planet. Sci. Lett.* **236**, 167–181 (2005).
- Mott, N. F. & Gurney, R. W. *Electronic Processes in Ionic Crystals* (Oxford Univ. Press, Oxford, 1948).
- Wood, B. J., Btyndzia, L. T. & Johnson, K. E. Mantle oxidation state and its relationship to tectonic environment and fluid speciation. *Science* **248**, 337–345 (1990).
- Williams, H. M. et al. Iron isotope fractionation and the oxygen fugacity of the mantle. *Science* **304**, 1656–1659 (2004).
- Utada, H., Koyama, T., Shimizu, H. & Chave, A. D. A semi-global reference model for electrical conductivity in the mid-mantle beneath the north Pacific region. *Geophys. Res. Lett.* **30**, 1194, doi:10.1029/2002GL016902 (2003).
- Lizzaralde, D., Chave, A. D., Hirth, G. & Schultz, A. Northeastern Pacific mantle conductivity profile from long-period magnetotelluric sounding using Hawaii to California submarine cable data. *J. Geophys. Res.* **100**, 17837–17854 (1995).
- The MELT seismic team. Imaging the deep seismic structure beneath a mid-ocean ridge: The MELT experiment. *Science* **280**, 1215–1218 (1998).
- Jung, H. & Karato, S. Water-induced fabric transitions in olivine. *Science* **293**, 1460–1463 (2001).
- Presnall, D. C., Simmons, C. L. & Porath, H. Change of electrical conductivity of a synthetic basalt during melting. *J. Geophys. Res.* **77**, 5665–5672 (1972).
- Waff, H. S. & Bulau, J. R. Equilibrium fluid distribution in an ultramafic partial melt under hydrostatic conditions. *J. Geophys. Res.* **84**, 6109–6114 (1979).
- Roberts, J. J. & Tyburczy, J. A. Partial-melt electrical conductivity: Influence of melt composition. *J. Geophys. Res.* **104**, 7055–7066 (1999).

24. Takahashi, E. Melting of a dry peridotite KLB-1 up to 14 GPa: Implications on the origin of peridotitic upper mantle. *J. Geophys. Res.* **91**, 9367–9382 (1986).
25. Kushiro, I. *et al.* Melting of a peridotite nodule at high pressures and high water pressures. *J. Geophys. Res.* **73**, 6023–6029 (1968).
26. Wyllie, P. J. & Huang, W.-I. Influence of mantle CO₂ in the generation of carbonatites and kimberlites. *Nature* **257**, 297–299 (1975).
27. Wang, W. & Takahashi, E. Subsolidus and melting experiments of K-doped peridotite KLB-1 to 27 GPa: Its geophysical and geochemical implications. *J. Geophys. Res.* **105**, 2855–2868 (2000).
28. Holtzman, B. K. *et al.* Melt segregation and strain partitioning: implications for seismic anisotropy and mantle flow. *Science* **301**, 1227–1230 (2003).
29. Baba, K., Chave, A. D., Evans, R. L., Hirth, G. & Mckie, R. L. Mantle dynamics beneath the East Pacific Rise at 17°S: Insights from the Mantle Electromagnetic and Tomography (MELT) experiment. *J. Geophys. Res.* **111**, B02101, doi:10.1029/2004JB003598 (2006).
30. Dunn, R. A. & Forsyth, D. W. Imaging the transition between the region of mantle melt generation and the crustal magma chamber with short-period Love wave propagation along the southern East Pacific Rise. *J. Geophys. Res.* **108**, B72352, doi:10.1029/2002JB002217 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank K. Baba, K. Fujita, M. Ichiki, T. Koyama, E. Ito and D. Yamazaki for discussions, D. Kohlstedt for comments and suggestions that improved manuscripts, and C. Oka for technical assistance.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to T.Y. (tyoshino@misasa.okayama-u.ac.jp).

The effect of water on the electrical conductivity of olivine

Duojun Wang^{1,2,3}, Mainak Mookherjee³, Yousheng Xu^{3,4} & Shun-ichiro Karato³

It is well known that water (as a source of hydrogen) affects the physical and chemical properties of minerals—for example, plastic deformation^{1–3} and melting temperature⁴—and accordingly plays an important role in the dynamics and geochemical evolution of the Earth. Estimating the water content of the Earth's mantle by direct sampling provides only a limited data set from shallow regions (<200 km depth)⁵. Geophysical observations such as electrical conductivity are considered to be sensitive to water content⁶, but there has been no experimental study to determine the effect of water on the electrical conductivity of olivine, the most abundant mineral in the Earth's mantle. Here we report a laboratory study of the dependence of the electrical conductivity of olivine aggregates on water content at high temperature and pressure. The electrical conductivity of synthetic polycrystalline olivine was determined from a.c. impedance measurements at a pressure of 4 GPa for a temperature range of 873–1,273 K for water contents of 0.01–0.08 wt%. The results show that the electrical conductivity is strongly dependent on water content but depends only modestly on temperature. The water content dependence of conductivity is best explained by a model in which electrical conduction is due to the motion of free protons. A comparison of the laboratory data with geophysical observations^{7–10} suggests that the typical oceanic asthenosphere contains $\sim 10^{-2}$ wt% water, whereas the water content in the continental upper mantle is less than $\sim 10^{-3}$ wt%.

Karato⁶ proposed that hydrogen enhances electrical conductivity in olivine either directly through charge transport by protons or indirectly through enhancement of ionic conductivity by Mg-Fe diffusion. Although this hypothesis has been used to interpret results of geophysical measurements of electrical conductivity^{10–13}, there has been no experimental data to test this hypothesis. In fact, the recent study of wadsleyite and ringwoodite¹⁴ does not entirely support Karato's hypothesis: electrical conduction in these minerals (with hydrogen) occurs through the motion of free protons—but not of all protons as Karato⁶ assumed (protons bound in the M-sites do not contribute to conductivity directly). If this is the case for olivine, then the results of these previous studies will be invalidated, and the issue of estimating water content from electrical conductivity would have to be revisited. Consequently, we have conducted an experimental study of the influence of hydrogen on electrical conductivity in olivine, and applied the results to infer the water (hydrogen) content of the upper mantle. In this Letter we report the first results and present discussions of their implications.

The starting material was San Carlos olivine. Inclusion-free, clear crystals were selected, and crushed into fine grains. Grains with different sizes were sorted by the sedimentation method. We added $\sim 2\%$ orthopyroxene to control the silica activity. A powder sample is packed in a nickel capsule and hot-pressed at ~ 4 GPa, 1,373–1,473 K

for ~ 1 – 3 h using a Kawai-type multianvil press (the presence of NiO was confirmed after the synthesis). A small amount of water was added by the decomposition of a talc+brucite mixture. Water contents of samples were measured both before and after each measurement of electrical conductivity by Fourier-transform infrared (FT-IR) spectroscopy. The calibration given in ref. 15 was used to calculate the water content of olivine grains from infrared absorption; corrections were made for the contribution of 'grain-boundary' water (see Methods). Grain sizes of hot-pressed samples are ~ 10 – 30 μm . The distribution of crystallographic orientations is nearly random, as determined by electron backscatter diffraction measurements.

Figure 1 shows the sample assembly for electrical conductivity measurements. A disk-shaped sample (~ 1.6 mm diameter, ~ 0.4 mm thickness) is placed between two Ni electrodes that are surrounded by alumina rings. All metallic components in the sample assembly are made of Ni to make sure that the oxygen fugacity during the measurements is close to that of the Ni–NiO buffer at which

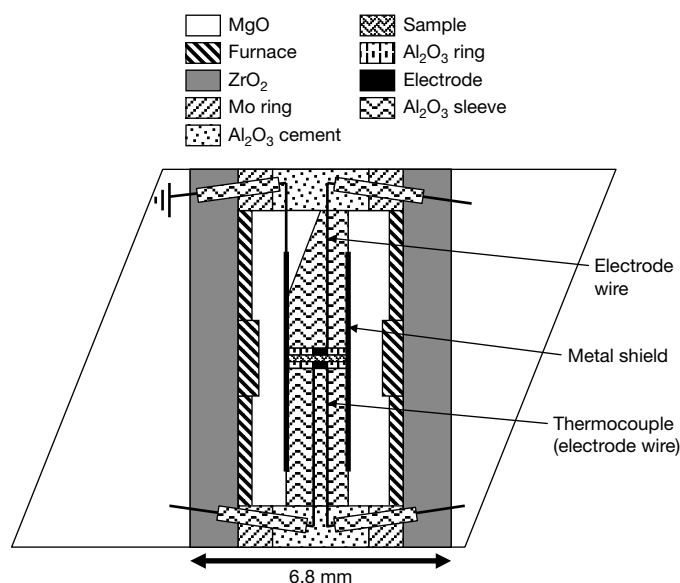


Figure 1 | The sample assembly for electrical conductivity measurements at high pressure. A polycrystalline disk-shaped sample (diameter 1.6 mm, thickness 0.4 mm) is inserted between two alumina rods that contain Ni electrodes. Note that a thermocouple is in contact with the sample so that the temperature at which conductivity is measured can be determined with a small uncertainty. The response of the sample to an a.c. voltage is recorded as a function of frequency to obtain the complex impedance. The resistance is calculated from the complex impedance. The sample space is separated from the furnace by a thin cylinder of Ni that acts as a shield to minimize the noise and the leak current.

¹Institute of Geochemistry, Chinese Academy of Sciences, Guiyang 550002, China. ²Institute of Geology, China Earthquake Administration, Beijing 100029, China. ³Department of Geology and Geophysics, Yale University, New Haven, Connecticut 06520, USA. ⁴Department of Mathematics, University of Connecticut, Storrs, Connecticut 06269, USA.

hot-pressed specimens were equilibrated. A cylindrical Ni foil is placed between the sample and the furnace to minimize the leakage currents (for details, see ref. 16). The temperature was measured by a W5%Re–W26%Re thermocouple that is placed next to the sample and is also used as one of the electrodes. The temperature reading is stable better than ~ 10 K during one measurement. The complex impedances of samples were determined by a Solartron 1260 impedance/gain phase analyser at 4 GPa and 873–1,273 K within the frequency range 10^3 – 10^6 Hz with a voltage of 0.1 V. The choice of

relatively low voltage and a high frequency range was made to minimize the hydrogen loss during an experiment. All the data shown here are from measurements in which water loss (calculated from water content before and after the experiment) is less than $\sim 20\%$. A complete list of the results is provided in a Table in Supplementary Information 1. The conductivity values were calculated from the complex impedance of samples. The results are shown in Fig. 2. The error in the conductivity measurement is largely due to the resistance fit and are $\pm 5\%$, the errors in temperature measurements are ± 10 K or better, and the major errors are those for water content and are $\pm 20\%$ (see Methods).

Electrical conductivities of a mineral containing both hydrogen and iron can be written as:

$$\sigma = \sigma_{\text{Fe}} + \sigma_{\text{H}} \quad (1)$$

where σ_{Fe} is the electrical conductivity due to ferric iron, and σ_{H} is the electrical conductivity due to hydrogen. σ_{H} may be written as:

$$\sigma_{\text{H}} = AC_{\text{w}}^r \exp\left(-\frac{H^*}{RT}\right) \quad (2)$$

where we used a relation that the electrical conductivity is proportional to the concentration of charged species and their mobility (Nernst–Einstein relation), and the concentration of charged species is proportional to C_{w}^r , where C_{w} is the water content, A and r are constants, H^* is the activation enthalpy, R is the gas constant and T is temperature. The results are summarized in Table 1 together with the results for ‘dry’ olivine. The errors in estimated parameters are for one standard deviation, and errors in temperature, conductivity and water content measurements are included in the error estimate of these parameters. The range of grain sizes is too narrow to constrain any grain-size sensitivity, but a comparison of the present results with the results on a coarse-grained dunite¹⁷ (~ 1 mm grain size) suggests that the grain-size effect, if any, is small.

For hydrogen-free (‘dry’) olivine, electrical conductivity (σ_{dry}) is due mainly to the migration of electron holes that are created by ferric iron and $\sigma_{\text{dry}} \approx \sigma_{\text{Fe}}$ (refs 18, 19; note, however, that σ_{Fe} is in general dependent on hydrogen (water) content and $\sigma_{\text{dry}} = \lim_{C_{\text{w}} \rightarrow 0} \sigma_{\text{Fe}}$). The values of electrical conductivity in our hydrogen-bearing olivine are significantly higher than those of hydrogen-free (‘dry’) olivine²⁰ (Fig. 2a). Also, the activation enthalpy for ‘dry’ olivine (~ 154 kJ mol^{−1}) is much higher than that of ‘wet’ olivine ($\sim 87 \pm 5$ kJ mol^{−1}), and there are very small pressure effects on electrical conductivity²¹. Therefore, the conduction mechanism is different between ‘wet’ and ‘dry’ samples. For dry olivine samples, the conduction mechanism is considered to be the migration of electron holes^{18,19}. However, for ‘wet’ olivine samples, it should be due to a different charged species, hydrogen: $\sigma_{\text{wet}} \approx \sigma_{\text{H}}$.

Let us examine how hydrogen enhances electrical conductivity. Karato⁶ assumed that the electrical conductivity in olivine containing hydrogen is dominated by the diffusion of hydrogen, and used the experimental data on hydrogen diffusion and solubility to calculate the electrical conductivity. The data that Karato⁶ used were chemical diffusion of hydrogen, but a later analysis by Kohlstedt and Mackwell²² showed that the data on chemical diffusion roughly correspond to self diffusion of protons (corrected by a factor of 2). Consequently, we use the data by Kohlstedt and Mackwell²² and

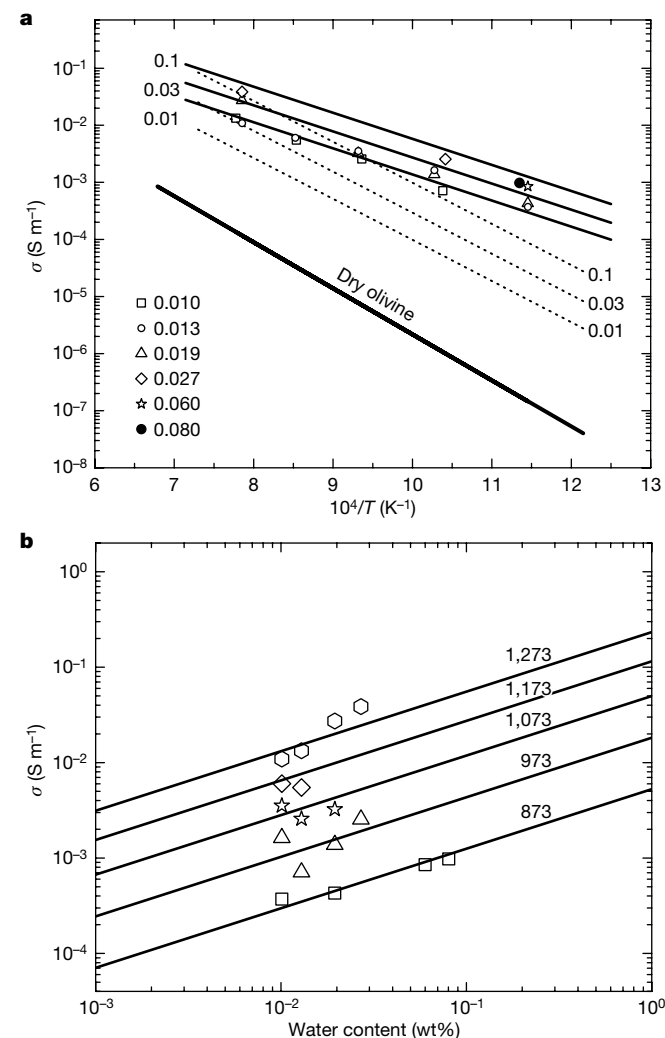


Figure 2 | Electrical conductivity versus inverse temperature and water content. **a**, A plot of electrical conductivity (σ) versus inverse temperature. The numbers next to each line indicate the water content (in wt%). Solid lines with numbers are the results of multilinear regression from all the data. Each symbol represents the data corresponding to a given water content (in wt%). The broken lines show the conductivity values calculated from water diffusion coefficients determined by Kohlstedt and Mackwell²² using the model by Karato⁶, and the conductivity for ‘dry’ olivine²⁰. Both the magnitude and the slope from the present study are different from those of ‘dry’ olivine and the trend calculated using the Karato model, indicating that the charge carrier corresponding to the present study is different from those in the previous studies (neutral hydrogen-defect at M-sites for the Karato model, Fe³⁺ for ‘dry’ olivine). **b**, A plot of electrical conductivity versus water content. The fit of the data to a model equation yields $r = 0.62 \pm 0.15$. The numbers along each line represent the temperature in K. A relatively large error for r is due to the fact that only a narrow range of water content can be explored for olivine due to the relatively small hydrogen solubility, and to a large error in the estimate of water content from FT-IR (see Methods). Each symbol corresponds to the data for a given temperature (in K).

Table 1 | Parameter values for electrical conductivity of olivine

Olivine conditions	$\log_{10}[A \text{ (S m}^{-1}\text{)}]$	r	H^* (kJ mol ^{−1})
Wet	3.0 ± 0.4	0.62 ± 0.15	87 ± 5
Dry	2.4	0	154

The relation $\sigma = AC_{\text{w}}^r \exp(-H^*/RT)$ is used. Errors are for one standard deviation, and include the contribution of errors in individual measurements (error in temperature, water content and conductivity). The data for ‘dry’ conditions are from ref. 20. The electrical conductivity under ‘dry’ conditions corresponds to conductivity due to ferric iron at zero hydrogen content ($\sigma_{\text{dry}} = \lim_{C_{\text{w}} \rightarrow 0} \sigma_{\text{Fe}}$).

the Nernst–Einstein equation to calculate electrical conductivity in olivine. The conductivity of an aggregate is calculated from the single crystal diffusion data by taking a Hashin–Shtrikman average²³. The results are compared with our experimental data. We note that the absolute values of electrical conductivity calculated from the Karato⁶ model are significantly lower than the actually measured values at low temperatures, whereas the measured values are in good agreement with the Karato model at high temperatures. Also, the activation enthalpy of electrical conductivity ($87 \pm 5 \text{ kJ mol}^{-1}$) is significantly smaller than that of diffusion of hydrogen ($150 \pm 30 \text{ kJ mol}^{-1}$)²² (Methods). Therefore we conclude that the mechanism of electrical conductivity in olivine is different from Karato's model⁶.

Huang *et al.*¹⁴ reached a similar conclusion, and suggested electrical conduction by free protons produced by an ionization reaction:



where H'_{M} is a proton trapped at an M-site vacancy and $\text{H}^{\bullet}$ is a free proton. This mechanism predicts $r = 0.75$ (for the charge neutrality condition of $[\text{H}'_{\text{M}}] = [\text{Fe}^{\bullet}_{\text{M}}]$) or 0.50 (for $2[\text{V}''_{\text{M}}] = [\text{Fe}^{\bullet}_{\text{M}}]$ where V''_{M} is a vacancy at an M-site) as opposed to $r = 1$ for the Karato model⁶. The value of r in our study is 0.62 ± 0.15 (Fig. 2b), which is significantly lower than the Karato model⁶, and is consistent with the free proton model, although the errors in r are admittedly large (due mainly to the large errors in water content measurements).

An alternative mechanism by which hydrogen may enhance electrical conductivity is through the enhancement of diffusion of Mg (Fe). However, a recent experimental study²⁴ indicates that the influence of conduction by diffusion of Mg (Fe) is small.

The chemical conditions of our experiments (that is, oxygen fugacity and oxide activity) are nearly identical to those in the Earth's upper mantle. The pressure (4 GPa) at which electrical conductivity was measured corresponds to the typical upper asthenosphere ($\sim 120 \text{ km}$ depth). However, the temperature range used in our experiments ($873\text{--}1,273 \text{ K}$) is lower than a typical temperature of

the asthenosphere ($\sim 1,600\text{--}1,700 \text{ K}$). Therefore we extrapolate our results in terms of temperature and compare the results with some of the geophysical observations of electrical conductivity in the upper mantle ($\sim 100\text{--}200 \text{ km}$) (Fig. 3). This extrapolation causes only small uncertainties because the activation enthalpy is well constrained by the present study (see Table 1 and Methods) and the extrapolation in $1/T$ space is not large compared to the range of $1/T$ space in which the measurements were made. The electrical conductivity of the oceanic asthenosphere is typically $\sim 10^{-1} \text{ S m}^{-1}$ (refs 9–11), which corresponds to a water content of $\sim 0.8 \times 10^{-2} \text{ wt\%}$ ($\sim 1,500 \text{ p.p.m. H/Si}$) (error is $\sim \pm 50\%$). This value agrees well with the estimated water content from the composition of mid-ocean ridge basalt (see ref. 25 for a review). However, geophysical observations^{7–11} also show that electrical conductivity has a large regional variation. The electrical conductivity in the continental upper mantle is typically lower, on the order of $\sim 10^{-2} \text{ S m}^{-1}$ (refs 7, 8). These values correspond to a significantly lower water content, on the order of $\sim 10^{-3} \text{ wt\%}$ or lower.

We also note that if hydrogen is indeed responsible for electrical conductivity in hydrogen-rich olivine, and if electrical conduction occurs mostly through grains, then one would expect anisotropy in electrical conductivity if olivine assumes strong lattice-preferred orientation. Some geophysical observations suggest anisotropy in electrical conductivity in the upper mantle^{26,27}, but a quantitative assessment of anisotropy in electrical conductivity will require an experimental study on strongly textured olivine aggregates.

We conclude that our experimental study has established a solid foundation for the use of geophysically inferred electrical conductivity to infer water content (and temperature) in the upper mantle of the Earth. Using observations of geomagnetic field from networks including ocean-bottom electro-magnetometers, high-resolution electrical conductivity profiles in three dimensions are becoming available. A combination of these results with laboratory studies will provide important data to constrain the distribution of hydrogen in the upper mantle of the Earth, which will provide a clue to the evolution of this region of our planet.

METHODS

Determination of water content. The water (hydrogen) content of samples was measured on polycrystalline samples using FT-IR absorption spectroscopy. An IR absorption by a polycrystalline sample is caused by OH-related species in grains as well as these species on grain boundaries. Our IR absorption spectra contain both a broad peak centred at $\sim 3,400 \text{ cm}^{-1}$ and sharp peaks at $\sim 3,500\text{--}3,600 \text{ cm}^{-1}$. Previous results suggest that a broad peak observed in polycrystalline samples is due to water on grain boundaries³ or to quenched melt²⁸. Consequently, we subtracted a broad peak at $\sim 3,400 \text{ cm}^{-1}$ from the original spectrum to estimate the water content in the grains in each sample (see Supplementary Information 2). The water content in the grains in a polycrystalline sample estimated in this way agrees reasonably well with the solubility of water reported in ref. 29 (with an error of $\pm 20\%$).

Activation enthalpy. The activation enthalpy determined in our experiments includes the temperature dependence of the concentration of defects ($\propto \exp(-H_f^*/RT)$, where H_f^* is the enthalpy for defect formation) and the temperature dependence of the mobility of defects ($\propto \exp(-H_m^*/RT)$, where H_m^* is the enthalpy for defect mobility). Therefore the activation enthalpy is $H^* = H_f^* + H_m^*$. In our experiments, the total water content in a sample in a series of experiments at various temperatures is kept nearly constant. However, the fraction of charge carrying species, $\text{H}^{\bullet}$, may change with temperature (in other words, H_f^* in our case corresponds to the enthalpy change associated with the reaction given in equation (3)). Such is probably the case in the Earth's mantle where minerals are under-saturated with water and the system is nearly closed. Consequently, the application of our data on the activation enthalpy to the Earth's interior is probably justified.

Note, however, that the total hydrogen content in diffusion experiments changes with temperature. The difference in the activation enthalpy between diffusion experiments ($150 \pm 30 \text{ kJ mol}^{-1}$) and electrical conductivity ($87 \pm 5 \text{ kJ mol}^{-1}$) could be due to the temperature dependence of hydrogen solubility, which is proportional to $\exp(-H_{\text{sol}}^*/RT)$ with $H_{\text{sol}}^* \approx 50 \text{ kJ mol}^{-1}$.

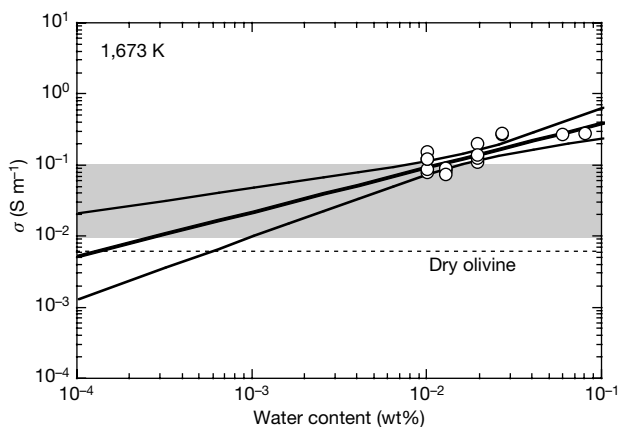


Figure 3 | Comparison of the conductivity of hydrogen-bearing olivine with geophysically inferred electrical conductivity for the asthenosphere.

Symbols show the calculated conductivity at $T = 1,673 \text{ K}$ from our results. The thick line indicates a best-fit line for the data, and the two thinner curves show the error envelope corresponding to one standard deviation. The electrical conductivity of the oceanic asthenosphere is typically $\sim 10^{-1} \text{ S m}^{-1}$ (refs 10, 11, 30), whereas that in the continental deep upper mantle (in the same depth range) is $\sim 10^{-2} \text{ S m}^{-1}$ (refs 7, 8). Shown also (horizontal broken line) is the electrical conductivity of hydrogen-free olivine. The thick horizontal grey bar indicates a range of typical values of electrical conductivity of the upper mantle ($\sim 100\text{--}200 \text{ km}$). We conclude that for the oceanic asthenosphere, the water content is estimated to be $\sim 0.8 \times 10^{-2} \text{ wt\%}$ ($\sim 1,500 \text{ p.p.m. H/Si}$), which is in excellent agreement with geochemical inference²⁵. In contrast, the water content in the continental upper mantle is significantly lower, and the conductivity values there are consistent with those of 'dry' olivine.

Received 31 March; accepted 11 September 2006.

- Karato, S. & Jung, H. Effects of pressure on high-temperature dislocation creep in olivine polycrystals. *Phil. Mag. A* **83**, 401–414 (2003).
- Karato, S., Paterson, M. S. & Fitz Gerald, J. D. Rheology of synthetic olivine aggregates: influence of grain-size and water. *J. Geophys. Res.* **91**, 8151–8176 (1986).
- Mei, S. & Kohlstedt, D. L. Influence of water on plastic deformation of olivine aggregates, 1. Diffusion creep regime. *J. Geophys. Res.* **105**, 21457–21469 (2000).
- Kushiro, I., Syono, Y. & Akimoto, S. Melting of a peridotite nodule at high pressures and high water pressures. *J. Geophys. Res.* **73**, 6023–6029 (1968).
- Bell, D. R. & Rossman, G. R. Water in Earth's mantle: The role of nominally anhydrous minerals. *Science* **255**, 1391–1397 (1992).
- Karato, S. The role of hydrogen in the electrical conductivity of the upper mantle. *Nature* **347**, 272–273 (1990).
- Olsen, N. Long-period (30 days - 1 year) electromagnetic sounding and the electrical conductivity of the lower mantle beneath Europe. *Geophys. J. Int.* **138**, 179–187 (1999).
- Tarits, P., Hautot, S. & Perrier, F. Water in the mantle: Results from electrical conductivity beneath the French Alps. *Geophys. Res. Lett.* **31**, doi:10.1029/2003GL019277 (2004).
- Evans, R. L. et al. Geophysical evidence from the MELT area for compositional control on oceanic plates. *Nature* **437**, 249–252 (2005).
- Lizarralde, D., Chave, A., Hirth, G. & Schultz, A. Northeastern Pacific mantle conductivity profile from long-period magnetotelluric sounding using Hawaii-to-California cable data. *J. Geophys. Res.* **100**, 17837–17854 (1995).
- Baba, K., Chave, A. D., Evans, R. L., Hirth, G. & Randall, L. Mantle dynamics beneath the East Pacific Rise at 17°S: Insights from the Mantle Electromagnetic and Tomography (MELT) experiments. *J. Geophys. Res.* **111**, doi:10.1029/2004JB003598 (2006).
- Simpson, F. & Tommasi, A. Hydrogen diffusivity and electrical anisotropy of a peridotite mantle. *Geophys. J. Int.* **160**, 1092–1102 (2005).
- Hirth, G., Evans, R. L. & Chave, A. D. Comparison of continental and oceanic mantle electrical conductivity: Is Archean lithosphere dry? *Geochim. Geophys. Res.* **5**, doi:10.1029/2000GC000048 (2000).
- Huang, X., Xu, Y. & Karato, S. Water content of the mantle transition zone from the electrical conductivity of wadsleyite and ringwoodite. *Nature* **434**, 746–749 (2005).
- Paterson, M. S. The determination of hydroxyl by infrared absorption in quartz, silicate glass and similar materials. *Bull. Mineral.* **105**, 20–29 (1982).
- Xu, Y., Poe, B. T., Shankland, T. J. & Rubie, D. C. Electrical conductivity of olivine, wadsleyite, and ringwoodite under upper-mantle conditions. *Science* **280**, 1415–1418 (1998).
- Wang, D. *Studies on Electrical Properties of Minerals and Rocks Under Defined Thermodynamic Conditions*. Ph.D. thesis, Chinese Acad. Sci. (2004).
- Karato, S. *Lattice Defects and Transport Properties of Olivine*. M.Sc. thesis, Univ. Tokyo (1974).
- Schock, R. N., Duba, A. G. & Shankland, T. J. Electrical conduction in olivine. *J. Geophys. Res.* **94**, 5829–5839 (1989).
- Constable, S., Shankland, T. G. & Duba, A. The electrical conductivity of an isotropic olivine mantle. *J. Geophys. Res.* **97**, 3397–3404 (1992).
- Xu, Y., Shankland, T. J. & Duba, A. G. Pressure effect on electrical conductivity of mantle olivine. *Phys. Earth Planet. Inter.* **118**, 149–161 (2000).
- Kohlstedt, D. L. & Mackwell, S. J. Diffusion of hydrogen and intrinsic point defects in olivine. *Z. Phys. Chem.* **207**, 147–162 (1998).
- Hashin, Z. & Shtrikman, S. A variational approach to the theory of effective magnetic permeability of multiphase materials. *J. Appl. Phys.* **33**, 3125–3131 (1962).
- Hier-Majumder, S., Anderson, I. M. & Kohlstedt, D. L. Influence of protons on Fe-Mg interdiffusion in olivine. *J. Geophys. Res.* **110**, doi:10.1029/2004JB003292 (2005).
- Bolfan-Casanova, N. Water in the Earth's mantle. *Mineral. Mag.* **69**, 229–257 (2005).
- Simpson, F. Intensity and direction of lattice-preferred orientation of olivine: are electrical and seismic anisotropies of the Australian mantle reconcilable? *Earth Planet. Sci. Lett.* **203**, 535–547 (2002).
- Gatzemeier, A. & Moorkamp, M. 3D modelling of electrical anisotropy from electromagnetic array data: hypothesis testing for different upper mantle conduction mechanisms. *Phys. Earth Planet. Inter.* **149**, 225–242 (2005).
- Stolper, E. M. Speciation of water in silicate melts. *Geochim. Cosmochim. Acta* **46**, 2609–2620 (1982).
- Kohlstedt, D. L., Keppler, H. & Rubie, D. C. Solubility of water in the α , β and γ phases of $(\text{Mg,Fe})_2\text{SiO}_4$. *Contrib. Mineral. Petrol.* **123**, 345–357 (1996).
- Ichiki, M. et al. Upper mantle conductivity structure of the back-arc region beneath northeastern China. *Geophys. Res. Lett.* **28**, 3773–3776 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements Z. Jing, Z. Jiang and I. Katayama provided the technical assistance that made this research possible. T. Kawazoe was most helpful with the error estimates. This work was supported by the NSF of China and the NSF of the United States.

Author Contributions S.-i.K. supervised the whole project and completed the paper. The experimental measurements of conductivity and the data analysis were made largely by D.W. and M.M. in collaboration with Y.X.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to S.-i.K. (shun-ichiro.karato@yale.edu).

A lamprey from the Devonian period of South Africa

Robert W. Gess¹, Michael I. Coates² & Bruce S. Rubidge¹

Lampreys are the most scientifically accessible of the remaining jawless vertebrates, but their evolutionary history is obscure. In contrast to the rich fossil record of armoured jawless fishes, all of which date from the Devonian period and earlier^{1–3}, only two Palaeozoic lampreys have been recorded, both from the Carboniferous period¹. In addition to these, the recent report of an exquisitely preserved Lower Cretaceous example⁴ demonstrates that anatomically modern lampreys were present by the late Mesozoic era. Here we report a marine/estuarine fossil lamprey from the Famennian (Late Devonian) of South Africa^{5,6}, the identity of which is established easily because many of the key specializations of modern forms are already in place. These specializations include the first evidence of a large oral disc, the first direct evidence of circumoral teeth and a well preserved branchial basket. This small agnathan, *Priscomyzon riniensis* gen. et sp. nov., is not only more conventionally lamprey-like than other Palaeozoic examples^{7,8}, but is also some 35 million years older. This finding is evidence that agnathans close to modern lampreys had evolved before the end of the Devonian period. In this light, lampreys as a whole appear all the more remarkable: ancient specialists that have persisted as such and survived a subsequent 360 million years.

Hyperoartia Müller, 1844

Petromyzontiformes Berg, 1940

Priscomyzon riniensis gen. et sp. nov.

Etymology. Generic name from Latin *prisco* (ancient) and *myzon* (a lamprey). Specific name from *Rini*, the Xhosa name for Grahamstown and the surrounding valley.

Holotype. Albany Museum, Grahamstown, Eastern Cape, catalogue number AM5750.

Locality and horizon. Waterloo Farm, Grahamstown, South Africa; Witpoort Formation, Witteberg Group, Famennian, Late Devonian^{5,6}.

Diagnosis. A small lamprey differing from all other living and fossil lampreys in having a large oral disc, the diameter of which approximately equals the length of the branchial region, and accounts for around one-half of total head length. A circumoral ring of 14 simple teeth is present, the posterior members of which are largest; there are no associated radiating series or plates of supplementary teeth. The first gill pouch is ventral to the otic capsule. Total body length is little more than 4.2 cm, post-branchial body shape is elongate and tapering, and only 1.5 times the head length. The dorsal fin extends from the immediate posterior of the branchial region.

Description. The 4.2-cm-long specimen of *Priscomyzon* is preserved in ventral view, in part (Fig. 1a) and counterpart (Fig. 1b) on a natural bedding plane. No post-depositional distortion is apparent in associated plant axes. The most striking feature of *Priscomyzon* is its large oral disc, edged with a soft outer lip, supported by an annular cartilage, and surrounding a circular mouth (Fig. 1c). This is the first clear evidence of a Palaeozoic lamprey with an oral disc, and the disc

is proportionately larger than those of similarly sized, post-metamorphic, living forms (Fig. 2). In contrast, the oral disc of the Late Carboniferous lamprey *Mayomyzon*^{7,8}, if present, is much smaller⁹, and no remnant of a disc is preserved in the Early Carboniferous lamprey *Hardistiella*¹⁰. The lamprey identity of the putative oral disc of *Pipiscus*¹¹ is uncertain¹².

Priscomyzon displays a set of 14 evenly spaced teeth surrounding the mouth. These are the first teeth to be discovered in any fossil lamprey, and resemble the circumoral arrangements of 19 or more teeth present in modern forms such as *Ichthyomyzon*, *Petromyzon*, *Caspiomyzon* and *Geotria*¹³. In *Priscomyzon*, the posterior circumoral teeth appear to be more elongate than the remainder, whereas in modern forms lateral or anterior teeth tend to be largest. Modern lamprey circumoral teeth usually display specialized shapes, and such teeth are often the largest of multiple series radiating to the oral disc perimeter. In comparison, the circumoral teeth of *Priscomyzon* are very simple, and in this respect probably primitive. Irregularly shaped objects in the centre of the mouth (oesophageal opening, Fig. 1c) might be traces of teeth from the apical cartilage of the 'tongue' complex, as in living hagfishes and lampreys¹⁴.

Several rod-like structures are preserved posterior to the oral disc. These include a possible styliform cartilage and parts of further cartilages from the underside of the neurocranium (Fig. 1). Ovoid patches flanking the midline are interpreted as the otic capsules. The lighter colour of these indicates denser mineralization, but otherwise the capsule material resembles that of surrounding skeletal remains. Otic capsules overlap the anterior of the branchial skeleton, a condition also seen in *Mayomyzon*^{7,8} but restricted to larval stages in living forms⁹ and absent from *Mesomyzon*⁴ (the reconstruction in Fig. 2a illustrates the difference in cranial layout). The position of the orbits is less clear because there are no darkened areas indicative of eye locations, as in *Mayomyzon*^{7,8}, *Hardistiella*¹⁰ and *Mesomyzon*⁴.

The branchial skeleton of *Priscomyzon* is preserved in greater detail than that of *Mesomyzon*⁴; gill arrangements in *Mayomyzon* are preserved only as dark imprints^{7,8}, whereas evidence of the branchial apparatus in *Hardistiella*¹⁰ is fragmentary¹⁵. In *Priscomyzon*, much of the right and parts of the left branchial baskets are preserved. The posterior five branchial arches are well defined, including evidence of at least two sets of horizontal bars (further bars may be obscured by matrix): the hypobranchials and either hypo- or epitrematic bars (Fig. 1). Anterior to these, areas of lesser mineral concentration suggest the presence of seven branchial pouches in total. Of the two Carboniferous lamprey species, *Mayomyzon* displays five pairs of gill pouches with indications of a further two⁷, and a possible juvenile specimen of *Hardistiella* shows at least six branchial units¹⁶. In *Priscomyzon* a bi-lobed structure posterior to the branchial skeleton corresponds to the expected position of the heart. However, it is not clear that this is evidence of a pericardial cartilage capsule, as in modern lampreys.

The post-branchial body of *Priscomyzon* is much narrower than the head and tapers posteriorly (Fig. 1). It is also exceptionally short:

¹Bernard Price Institute (Palaeontology), School for Geosciences, University of Witwatersrand, Johannesburg 2050, South Africa. ²Department of Organismal Biology, University of Chicago, Chicago, Illinois 60637, USA.

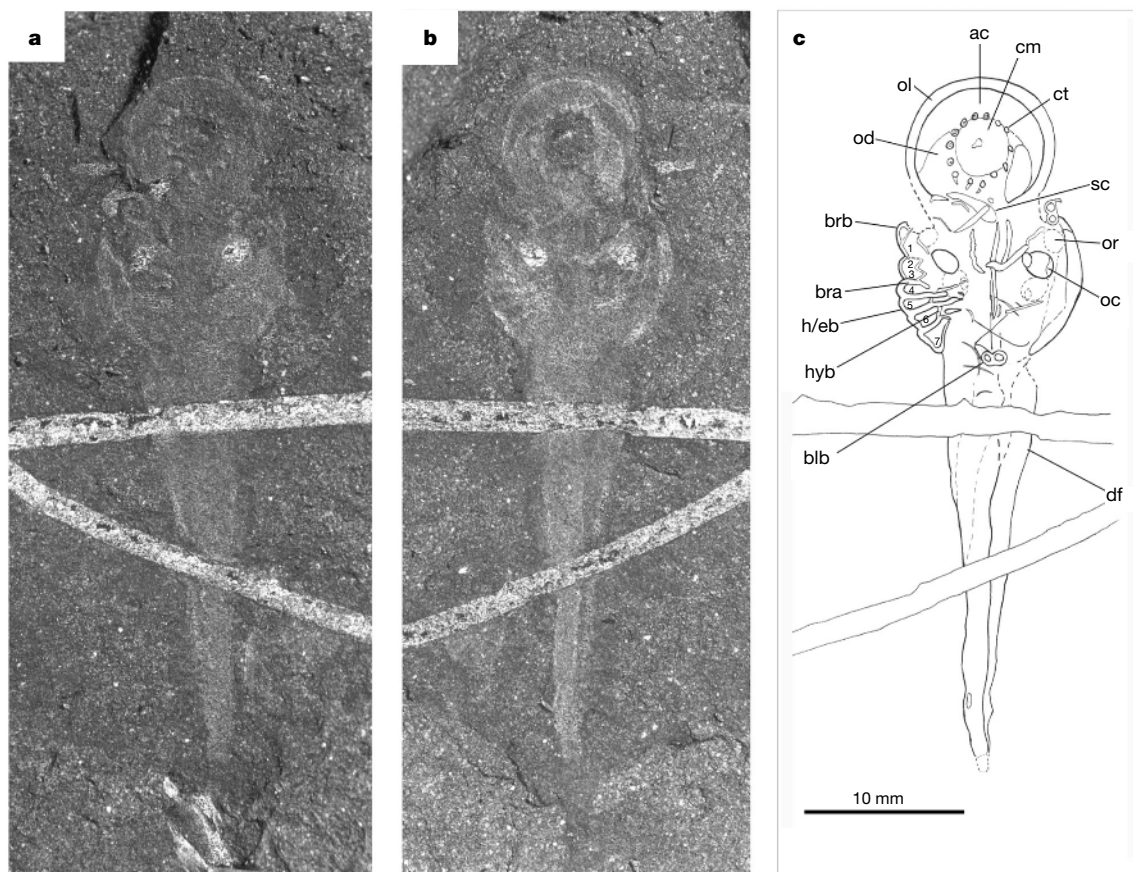


Figure 1 | Holotype of *Priscomyzon riniensis* gen. et sp. nov. This 360-million-year-old lamprey is the earliest example known in the fossil record, showing most of the specialized feeding structures present in modern forms. **a, b**, Part (**a**) and counterpart (**b**) of holotype AM5750. The total length of the specimen is 42 mm. **c**, Interpretive drawing of the holotype. ac, annular

cartilage; blb, bi-lobed structure; bra, branchial arch; brb, branchial basket; cm, circular mouth; ct, circumoral teeth; df, dorsal fin; hyb, hypobranchial bar; h/eb, hypotrematic/epitrematic bar; oc, otic capsule; od, oral disc; ol, outer lip; or, orbital region; sc, styliform cartilage; 1–7, positions of gill pouches.

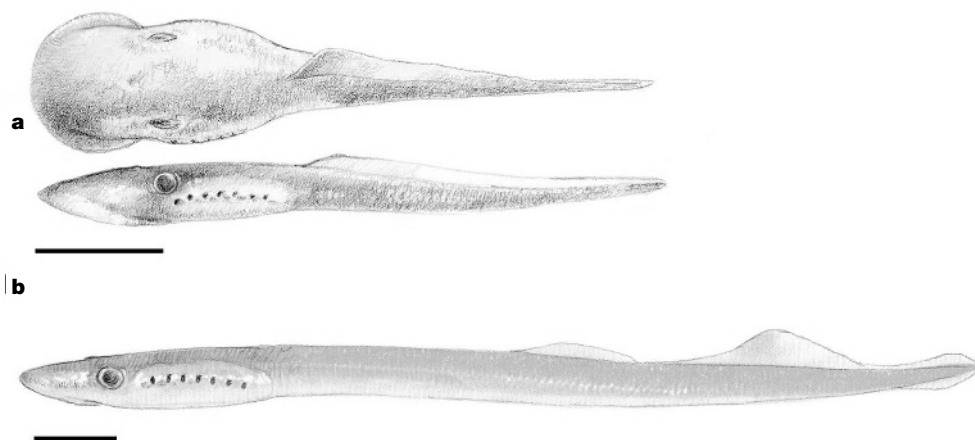


Figure 2 | Reconstruction of *Priscomyzon riniensis*, illustrating tadpole-like body proportions and large oral disc, compared with post-metamorphic modern lamprey, *Lampetra fluviatilis*. **a**, Reconstruction of *Priscomyzon* in dorsal (top) and left lateral (bottom) views. **b**, Macrophthalmia stage of *Lampetra*²⁸ showing anterior location of orbit and smaller oral disc, both

positioned in front of the branchial region. The total length of the specimen is 116 mm. Drawings in **a** and **b** are scaled to show equivalent head lengths: from anterior limit of the oral disc to rear of the branchial region. Horizontal bars indicate the anterior–posterior span of the oral disc in each species.

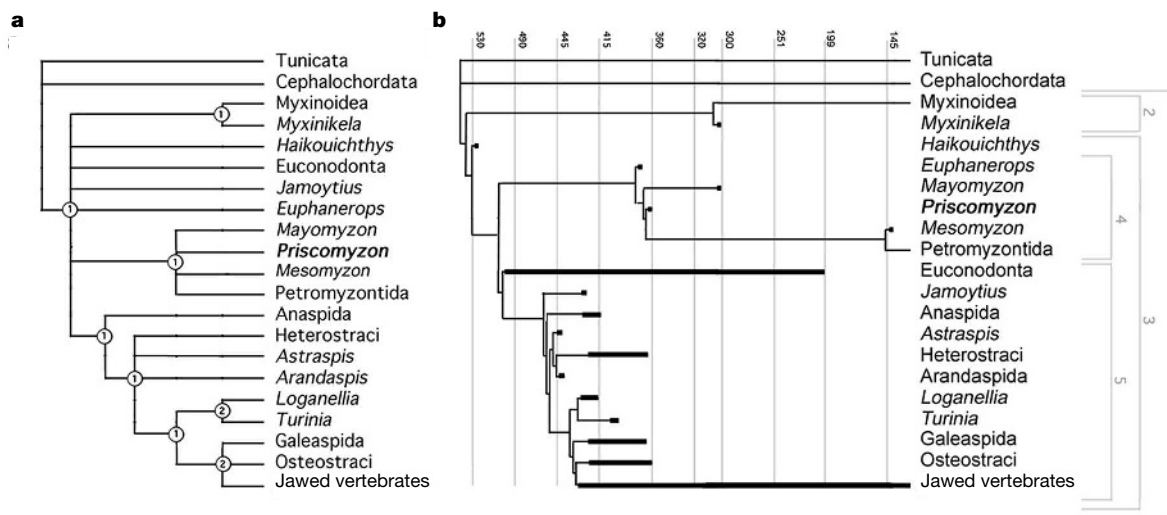


Figure 3 | Cladograms showing the hypothesized position of *Priscozoon* among early craniates. a, Strict consensus of 42 shortest trees: length, 211 steps; consistency index, 0.63; homoplasy index, 0.37; retention index, 0.7; rescaled consistency index, 0.44. *Priscozoon* lies within a polytomy of fossil and modern lampreys. Bremer support values are circled at nodes (compare with refs 2, 18; see Supplementary Information for details). **b**, Single tree

from analysis of re-weighted character set: cyclostomes (hagfishes and lampreys) are paraphyletic; *Euphanerops*^{1,3} is a stem lamprey on the basis of homoplastic synapomorphies. Time axis (million years) and temporal ranges of taxa (bold lines) are from refs 1, 3, 4, 12, 18, 23. Numbered brackets indicate major monophyletic groups: 1, Craniata; 2, Myxiniformes; 3, Vertebrata; 4, Petromyzontiformes; 5, Gnathostomata.

only 1.5 times the length of the head, compared to 3.3 times head length in the similarly sized *Mayomyzon*. Faint impressions of small lunate ridges are present, but their significance is unknown. Details of the dorsal fin are limited, although it is evident that it originates immediately behind the branchial region and extends as a continuous fold towards the caudal extremity. This resembles conditions in an ammocoete rather than in modern adult lampreys, in which separate anterior and posterior dorsal fins are located in the posterior half of the body (Fig. 2b). Like *Mayomyzon*, the single dorsal fin of *Priscozoon* is continuous with the caudal fin^{7,8}, whereas in *Hardistiella* the two fins might be more widely separated¹⁰.

To explore the phylogenetic position of *Priscozoon*, and examine the effects of these new data upon the existing hypotheses of relationships among jawless fossil and Recent vertebrates, we built upon data matrices from refs 2, 17 and 18. Relationships among these taxa are problematic because nucleotide sequence data tend to support cyclostome monophyly (hagfishes plus lampreys as sister group to gnathostomes)^{19,20}, whereas morphological analyses usually favour cyclostome paraphyly (hagfishes as a sister group to lampreys plus gnathostomes)^{2,18,20}. As well as adding new taxa⁴, our data set accounts for recent discoveries of lamprey-like conditions in putative stem gnathostomes^{3,21}, and includes characters describing similarities among the feeding apparatuses of living agnathans¹⁴.

Phylogenetic analysis²² of 21 taxa and 115 characters (see Supplementary Information) yields 42 shortest trees (most parsimonious solutions to the data set), with many of the major groupings found in ref. 2. A strict consensus of these places *Priscozoon* in a polytomy of fossil and Recent lampreys (Fig. 3a). A further polytomy at the base of vertebrates highlights increased instability among the relationships of 'unarmoured' fossil agnathans^{1,12,23}, although analysis of a reduced taxon set shows strong support for cyclostome paraphyly (see Supplementary Information). Analysis of the complete data set with enforced cyclostome monophyly (see Supplementary Information) increases tree length by only 8%, largely as a result of character losses along the hagfish branch, and with few changes to the gnathostome stem. Analysis of characters re-weighted after the first tree search (without enforced topological constraints) results in a single shortest tree (Fig. 3b), linking the 'naked anaspid'

*Euphanerops*³ to the base of a lamprey stem group. This result echoes previous suggestions about the relationships of these clades^{1,20}.

The discovery of *Priscozoon* within a Late Devonian marginal marine estuarine environment^{5,6} pushes the minimum date of lamprey-like fishes back by some 35 million years, and provides a new minimum date for molecular-clock-based estimates of the cyclostome crown node. The well developed oral disc, annular cartilages and circumoral teeth of *Priscozoon* suggests the evolutionary long-term stability of a highly specialized parasitic feeding habit. Lampreys have long been recognized as highly apomorphic¹³ but only now is it possible to appreciate just how ancient these specializations are. In this particular sense, lampreys might be described as 'living fossils'²⁴, and *Priscozoon* adds new phylogenetic perspective to studies using modern agnathans as model systems for deriving insight into primitive vertebrate conditions^{25–27}.

METHODS

Phylogenetic analysis was performed with the phylogenetic package PAUP*4.0b10 (ref. 22). See Supplementary Information for the list of 115 characters with sources of reference, the data matrix, and a strict consensus of six trees obtained when cyclostome monophyly was enforced as a topological constraint.

Received 1 June; accepted 7 August 2006.

1. Janvier, P. *Early Vertebrates* (Clarendon, Oxford, 1996).
2. Donoghue, P. C. J. & Smith, M. P. The anatomy of *Turinia pagei* (Powrie), and the phylogenetic status of the Thelodonti. *Trans. R. Soc. Edinb. Earth Sci.* **92**, 15–37 (2001).
3. Janvier, P., Desbiens, S., Willett, J. A. & Arsénault, M. Lamprey-like gills in a gnathostome-related Devonian jawless vertebrate. *Nature* **440**, 1183–1185 (2006).
4. Chang, M.-m. Zhang, J. & Miao, D. A lamprey from the Cretaceous Jehol biota of China. *Nature* **441**, 972–974 (2006).
5. Anderson, H. M., Hiller, N. & Gess, R. W. *Archaeopteris* (Progymnospermopsida) from the Devonian of southern Africa. *Bot. J. Linn. Soc.* **117**, 305–320 (1995).
6. Hiller, N. & Taylor, F. F. Late Devonian shore line changes: an analysis of Witteberg Group stratigraphy in the Grahamstown area. *South Afr. J. Geol.* **95**, 203–212 (1992).
7. Bardack, D. & Zangerl, R. First fossil lamprey: a record from the Pennsylvanian of Illinois. *Science* **162**, 1265–1267 (1968).
8. Bardack, D. & Zangerl, R. in *The Biology of Lampreys* Vol. 1 (eds Hardisty, M. W. & Potter, I. C.) 67–84 (Academic, London, 1971).

9. Hardisty, M. W. in *The Biology of Lampreys* (eds Hardisty, M. W. & Potter, I. C.), 3, 333–376 (Academic, London, 1981).
10. Janvier, P. & Lund, R. *Hardistiella montanensis* n. gen. et sp. (Petromyzontidae) from the Lower Carboniferous of Montana with remarks on the affinities of the Lampreys. *J. Vertebr. Paleontol.* **2**, 407–413 (1983).
11. Bardack, D. & Richardson, E. S. Jr. New agnathous fishes from the Pennsylvanian of Illinois. *Fieldiana Geology* **33**, 489–510 (1971).
12. Shu, D.-G. *et al.* Lower Cambrian vertebrates from south China. *Nature* **402**, 42–46 (1999).
13. Hubbs, C. L. & Potter, I. C. in *The Biology of Lampreys* Vol. 1 (eds Hardisty, M. W. & Potter, I. C.) 1–66 (Academic, London, 1971).
14. Yalden, D. Feeding mechanisms as evidence for cyclostome monophyly. *Zool. J. Linn. Soc.* **84**, 291–300 (1985).
15. Janvier, P. in *Recent Advances in the Origin and Early Radiation of Vertebrates* (eds Arratia, G., Wilson, V. H. & Schultze, H. P.) 29–52 (Dr Friedrich Pfeil, Munchen, 2004).
16. Lund, R. & Janvier, P. A second Lamprey from the Lower Carboniferous (Namurian) of Bear Gulch, Montana (U.S.A.). *Geobios* **19**, 647–652 (1986).
17. Janvier, P. The dawn of the vertebrates: characters versus common ascent in the rise of current vertebrate phylogenies. *Palaeontology* **39**, 259–287 (1996).
18. Donoghue, P. C. J., Forey, P. & Aldridge, R. J. Conodont affinity and chordate phylogeny. *Biol. Rev.* **75**, 191–251 (2000).
19. Delarbre, C., Gallut, C., Barriol, V., Janvier, P. & Gachelin, G. Complete mitochondrial DNA of the hagfish, *Eptatretus burgeri*: The comparative analysis of mitochondrial DNA sequences strongly supports the cyclostome monophyly. *Mol. Phylo. Evol.* **22**, 184–192 (2002).
20. Forey, P. & Janvier, P. Agnathans and the origin of jawed vertebrates. *Nature* **361**, 129–134 (1993).
21. Janvier, P. & Arsénault, M. Palaeobiology: Calcification of early vertebrate cartilage. *Nature* **417**, 609 (2002).
22. Swofford, D. L. *PAUP*: Phylogenetic analysis using parsimony (* and other methods), version 4.0b10* (Sinauer Associates, Sunderland, Massachusetts, 2003).
23. Shu, D.-G. *et al.* Head and backbone of the Early Cambrian vertebrate *Haikouichthys*. *Nature* **421**, 526–529 (2003).
24. Eldredge, N. in *Living Fossils* (eds Eldredge, N. & Stanley, S. M.) 272–277 (Springer, New York, 1984).
25. Cohn, M. J. Evolutionary biology: Lamprey *Hox* genes and the origin of jaws. *Nature* **416**, 386–387 (2002).
26. Takio, Y. *et al.* Evolutionary biology: Lamprey *Hox* genes and the evolution of jaws. *Nature* **429**, doi:10.1038/nature02616 (2004).
27. Pancer, Z. *et al.* Somatic diversification of variable lymphocyte receptors in the agnathan lamprey. *Nature* **430**, 174–180 (2004).
28. Hardisty, M. W. & Potter, I. C. in *The Biology of Lampreys* Vol. 1 (eds Hardisty, M. W. & Potter, I. C.) 127–206 (Academic, London, 1971).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful to B. de Klerk for his continuing support, and thank M. A. Purnell, R. J. Aldridge and M. M. Chang for advice and discussions. Work by R.W.G. and B.S.R. is supported by the Palaeontological Scientific Trust (PAST), National Research Foundation (NRF) and the Department of Science and Technology (DST) of South Africa, and that of M.I.C. by the Faculty Research Fund of the University of Chicago.

Author Contributions The discovery, identification, diagnosis and morphological description of the specimen are primarily the work of R.W.G. Phylogenetic analysis was performed by M.I.C., as well as final drafting of the paper. The research was supervised by B.S.R. All authors read, commented on and contributed to all parts of the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to R.G. (robg@imagineth.co.za).

Regulatory constraints in the evolution of the tetrapod limb anterior–posterior polarity

Basile Tarchini^{1†}, Denis Duboule¹ & Marie Kmita^{1†}

The anterior to posterior (A–P) polarity of the tetrapod limb is determined by the confined expression of *Sonic hedgehog* (*Shh*) at the posterior margin of developing early limb buds^{1,2}, under the control of HOX proteins encoded by gene members of both the *HoxA* and *HoxD* clusters^{3–6}. Here, we use a set of partial deletions to show that only the last four *Hox* paralogy groups can elicit this response: that is, precisely those genes whose expression is excluded from most anterior limb bud cells owing to their collinear transcriptional activation. We propose that the limb A–P polarity is produced as a collateral effect of *Hox* gene collinearity, a process highly constrained by its crucial importance during trunk development. In this view, the co-option of the trunk collinear mechanism, along with the emergence of limbs, imposed an A–P polarity to these structures as the most parsimonious solution. This in turn further contributed to stabilize the architecture and operational mode of this genetic system.

The exquisite functionalities of tetrapod limbs largely derive from their various polarities. Among them, the A–P asymmetry (from the thumb/radius to the minimus/ulna, respectively) results from the presence of an organizing centre, the zone of polarizing activity, at the posterior margin of the developing limb buds¹. The polarizing activity is achieved by the product of the gene *Sonic hedgehog* (*Shh*) (ref. 2), whose transcripts are found in posterior mesenchymal cells only, near the apical ectodermal ridge, from which growth factors such as *Fgf4* are necessary for *Shh* transcriptional onset and maintenance^{7–12}. *Shh* signalling is required to maintain the apical ectodermal ridge and hence to sustain the growth of the structure, as well as to organize limb patterns along the A–P axis^{13–16}. Therefore, the mechanism that restricts *Shh* transcription in posterior limb bud cells is critical for the generation of this polarity.

Gain-of-function experiments have suggested that some *Hox* gene products can trigger *Shh* transcription at ectopic sites of developing limb buds^{3–5}, probably through a direct interaction between HOX proteins, their cofactors and the *Shh* limb enhancer⁶, leading to modifications in A–P patterning. Recently, the deletion of both *HoxA* and *HoxD* clusters in developing forelimbs was reported to abrogate *Shh* expression, thus confirming that HOXA and/or HOXD proteins are required for *Shh* transcription¹⁷. However, several *Hoxa/d* genes are expressed throughout the emerging limb bud¹⁸, including in anterior cells, making the posterior restriction of *Shh* transcripts difficult to explain on this basis alone. Therefore, we set out to determine, in physiological conditions, which *Hox* product(s) could elicit *Shh* transcription in limb buds.

We generated a series of *Hoxa/Hoxd* double-mutant mice and started with a deletion of all *Hoxa* genes combined with the absence of the *Hoxd8* to *Hoxd13* interval. This resulted in the absence of *Shh* expression (Fig. 1b), even though the remaining *Hoxd1*, *Hoxd3* and *Hoxd4* genes were expressed throughout the limb mesenchyme at the

expected stage for *Shh* transcription¹⁸, suggesting that the latter gene products were unable to trigger *Shh* transcription. We next tested whether the addition of both *Hoxd8* and *Hoxd9* function could elicit *Shh* expression. There again, *Shh* transcripts were not detected (Fig. 1c). Interestingly, deleting from *Hoxd10* to *Hoxd13* led to *Hoxd9* upregulation in posterior cells (Fig. 1d), showing that even a robust

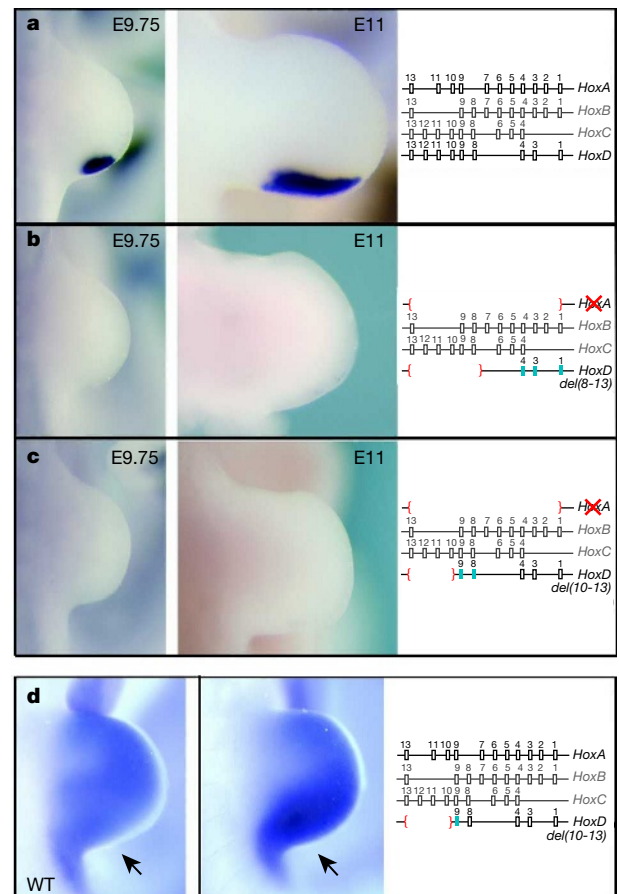


Figure 1 | Control of *Shh* expression by *Hoxa* and *Hoxd* genes. **a–c**, *Shh* expression in embryonic day (E) 9.75 and E11 limb buds, anterior to the top. The genotypes are depicted on the right-hand side. **a**, Wild type. **b**, *Shh* transcripts were absent from limb buds lacking all *Hoxa* genes and those from *Hoxd8* to *Hoxd13*. **c**, The addition of *Hoxd8* and *Hoxd9* functions did not elicit a *Shh* signal, even though *Hoxd9* was overexpressed posteriorly (**d**). *Hox* gene products from groups 1 to 9 are unable to activate and/or maintain *Shh* expression.

¹Department of Zoology and Animal Biology and National Research Centre 'Frontiers in Genetics', University of Geneva, Sciences III, Quai Ernest Ansermet 30, 1211 Geneva 4, Switzerland. [†]Present addresses: Department of Biology, McGill University, 1205 Avenue Docteur Penfield, Montreal H3A 1B1, Quebec, Canada (B.T.); Laboratory of Genetics and Development, Institut de Recherches Cliniques de Montréal (IRCM), Université de Montréal, 110 avenue des Pins Ouest, H2W 1R7, Montréal, Quebec, Canada (M.K.)

dose of *Hoxd9* delivered in the right cells was unable to trigger *Shh* expression (Fig. 1c).

We then used two additional combinations to re-introduce either *Hoxd10* or *Hoxd13* function to the former deletion (Fig. 2a). In both cases, *Shh* transcripts were detected unambiguously, indicating that either *Hoxd10* or *Hoxd13* products can elicit *Shh* transcription. However, neither of these two HOX products, in isolation, appeared to fully recapitulate the normal *Shh* pattern. *Shh* messenger RNAs were detected rather distally, in the presence of *Hoxd13* only (Fig. 2a, bottom), but more proximally in the presence of *Hoxd10* alone (Fig. 2a, top). Even though these differences may be slightly biased by distortions in the growth of the mutated limb buds themselves, these results suggested that the wild-type *Shh* spatial pattern is established in response to multiple HOX proteins. This was supported by limb buds lacking all *Hoxd* genes, as well as *Hoxa13*, the most distally expressed gene of the *HoxA* cluster, where *Shh* was expressed in a proximal domain only (Fig. 2b), probably in response to *Hoxa11* and/or *Hoxa10*. A quantitative aspect was also revealed by comparing *Shh* expression in *HoxA*^{+/+};*HoxD*^{del(8-13)/del(8-13)} versus *HoxA*^{+/-};*HoxD*^{del(8-13)/del(8-13)} early limb buds (Fig. 2c). A reduced *Shh* expression was associated with the removal of one dose of *Hoxa* genes. All together, this analysis showed that *Shh* expression is controlled both by quantitative and qualitative combinations of *Hoxa* and *Hoxd* genes, belonging to paralogous groups 10 to 13.

The corresponding phenotypic analyses indicated that forearm and hand development requires the function of those *Hoxa* and *Hoxd* genes that can trigger *Shh* expression (Fig. 3a). Forelimbs carrying *Hox* deletions preventing *Shh* activation were virtually identical in morphology to the full *HoxA/HoxD* deficiency¹⁷, with no genuine forearm and digits (Fig. 3b). Instead, addition of *Hoxd13* alone induced the formation of both digits and a remnant of forearm, whereas *Hoxd10* function induced a clear, yet truncated forearm followed by a series of rays (Fig. 3a). Therefore, the recruitment of the *Hox* system in growing limbs and its impact upon *Shh* activation led to both an increased number of bony elements along the A–P axis and their concomitant patterning, which illustrates a remarkable parsimony in producing appendages with high adaptive values. The resulting heteromeric nature of distal appendages was probably the ground for the architectural versatility and associated functional features that accompanied the emergence of tetrapods¹⁹. Also, while confirming the function of *Hoxd9* in proximal limb structures²⁰ (Fig. 3b), this analysis indicated that 3'-located genes (for example, *Hoxd4*, *Hoxd3*) have no function during limb development, despite their expression there. We discuss below (Fig. 4) why, paradoxically, these genes are nevertheless important in this context.

We believe the co-option of the collinear mechanism from the developing trunk to developing appendages is at the origin of the limb A–P polarity. In limb buds, *Hox* genes are activated in a spatial-temporal sequence following their genomic topography, leading to the progressive depletion of *Hox* gene transcripts from the anterior quadrant of the early bud, such that *Hoxd1* is transcribed throughout the bud, whereas *Hoxd13* is only expressed in a small posterior-distal domain¹⁸ (Fig. 4a). Interestingly, the transition between *Hox* genes expressed throughout the bud and those excluded from anterior cells is precisely between *Hoxd9* and *Hoxd10*, that is, between the last gene unable to activate *Shh* and the first one to achieve this task, respectively.

The existence of this transition (Fig. 4, arrowhead) is critical for the morphology of the limb. Should it disappear, or be shifted towards more 'posterior' (5') genes (for example, between *Hoxd12* and *Hoxd13*; Fig. 4b, arrowhead), expression of *Hox* genes restricted to the posterior mesenchyme would extend in anterior cells and trigger ectopic *Shh* transcription anteriorly. We would expect such limbs to be polydactylous with a bilateral symmetry³⁻⁵. Conversely, a premature occurrence of this transition would ultimately prevent *Shh* transcription owing to the lack of appropriate HOX products, and concurrent limb truncation (Fig. 4c). In this view, *Hox* gene

collinearity in the early limb bud is a critical parameter in organizing the A–P polarity. Such a mechanism must be well controlled to avoid variations that would override the buffering capacity of the system, and which would thus severely affect the final morphology. Expression of 3'-located genes (for example, *Hoxd1*, *Hoxd3*) in the early buds may help in this respect, despite the absence of morphological

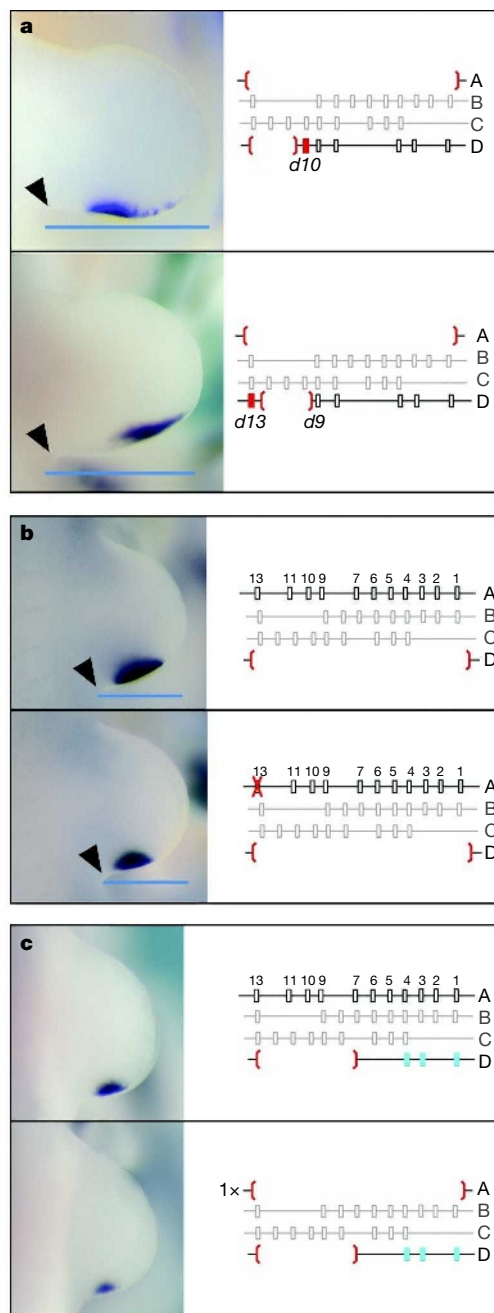


Figure 2 | 5'-located *Hox* genes induce *Shh* expression in a gene-specific and dose-dependent manner. **a**, The addition of either *Hoxd10* (top panel, red) or *Hoxd13* (bottom panel, red) function was enough to trigger *Shh* expression, yet with some qualitative differences; *Hoxd13* elicited *Shh* expression more distally than *Hoxd10* (arrowheads). **b**, Likewise, in the absence of *HoxD*, further removal of *Hoxa13* function seemed to restrict *Shh* expression proximally. **c**, A quantitative input was also observed, for example when comparing the potential of either two copies (top panel) or one (bottom panel) copy of *Hoxa* genes to activate *Shh* transcription, in the absence of all posterior *Hoxd* genes. Identical staining conditions showed a reduction in *Shh* transcripts (compare top and bottom panels). Blue lines in **a** and **b** indicate the proximal–distal extent of the limb bud at these stages.

impact, because even just their presence participates in delaying the activation of 5'-located genes (Fig. 4).

The early collinear mechanism acting in limb buds has recently been proposed to rely upon opposite regulations, located at both extremities of the *HoxD* cluster, one acting as a timer for gene activation, the other preventing expression in anterior cells¹⁸. Consequently, each gene of the cluster will show a unique response

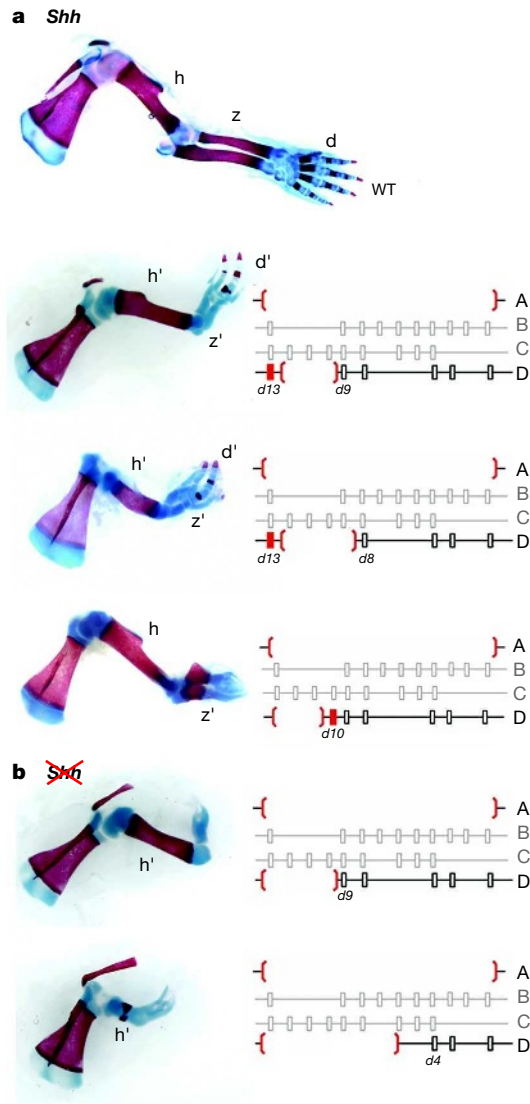


Figure 3 | Skeletons of the various mutant stocks in either the presence (top) or absence (bottom) of *Shh* signalling. **a**, On the top is a wild-type newborn forelimb. Below are mutant phenotypes. Genotypes are indicated, with all specimens carrying a conditionnal *HoxA* cluster deletion ($A^{c/-}$). As long as *Hoxd13* was present, both digits and a clear forearm (zeugopod; z, z') were observed, though ill-formed, reduced and poorly ossified. In contrast, the introduction of *Hoxd10* (bottom panel) led to a better-developed and ossified zeugopod, whereas only rod-like structures were present distally. The presence of an A–P polarity, as scored for example at the articulation between the humerus (h, h') and radius/ulna (z, z') coincides with the presence of *Shh* signalling. **b**, In contrast, when all posterior *Hoxd* genes were missing, neither digit nor radius/ulna were detected, just an abnormal and segmented cartilaginous condensation. These phenotypes were virtually identical to the full *HoxA/HoxD* cluster deletion¹⁷ and illustrated the absence of both *Shh* signalling and A–P polarity. The severity of the humerus truncation correlates with the number of remaining *Hox* genes, regardless of *Shh* activity, pointing to the function of groups 8 and 9 in the making of limb proximal structures (**b**, compare top and bottom panels). The presence, absence or orientation of the pectoral girdle results from the dissection of the specimen.

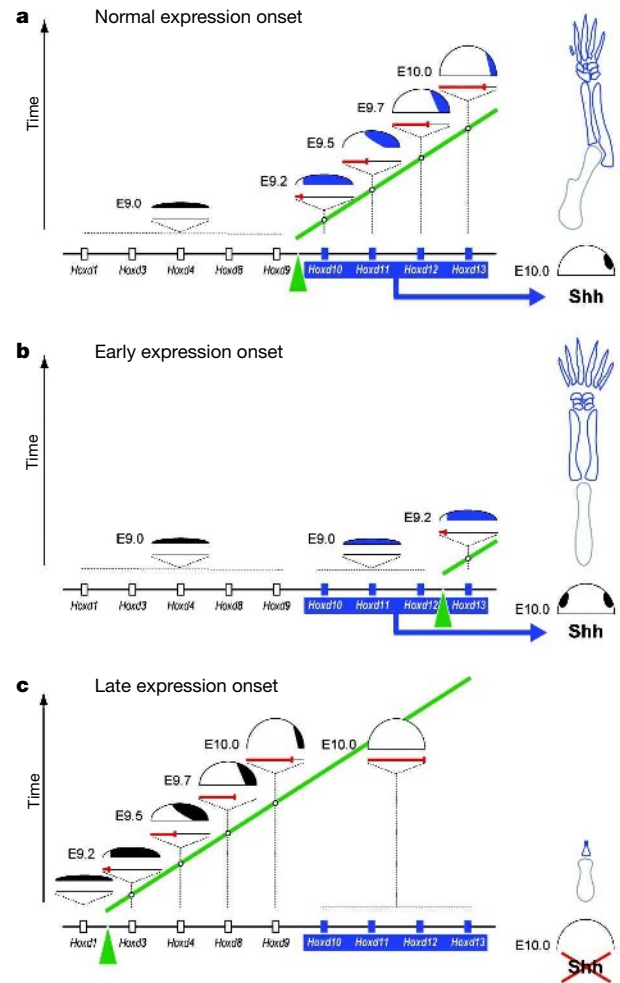


Figure 4 | Regulatory constraints in the emergence of the limb A–P polarity. This scheme shows how collinear *Hoxd* gene activation in early limb buds generates an A–P pattern and how temporal modifications (heterochronies) of this mechanism may affect limb morphology. **a**, Wild type. The expression of the different *Hoxd* genes is shown between E9.0 and E10.0, at transcriptional onset. The *HoxD* cluster is on the x axis and time on the y axis. Temporal collinearity thus appears as a positive slope (green line), posterior genes (for example, *Hoxd13*) being activated later than more anterior neighbours (for example, *Hoxd11*). Spatial collinearity, the progressive exclusion of *Hox* transcripts from the anterior-most region of the limb bud, is depicted with red bars. Genes and expression patterns in blue are those able to elicit *Shh* transcription, unlike those in black. Correspondingly, the skeletal elements are shown either in black or in blue, illustrating which *Hox* gene functions they require to properly develop. The beginning of *Hox* transcripts exclusion from anterior cells is indicated with a green arrowhead on the X axis. Anterior exclusion starts with *Hoxd10*, that is, precisely with the first HOX product able to trigger *Shh* transcription. Qualitative and quantitative inputs will lead to the activation of *Shh* at the posterior margin only, allowing for the formation of a fully grown, A–P polarized, limb skeleton. **b**, A substantial heterochronic shift of temporal collinearity towards premature *Hox* genes activation would lead to anterior exclusion only for the very last *Hox* gene transcripts (for example *Hoxd13*), with some 'blue' genes expressed throughout the early bud leading to *Shh* activation at both posterior and anterior margins and consequent bilateral symmetry of the limb skeleton. **c**, Conversely, a much delayed *Hox* gene activation would induce an anterior exclusion of 3' genes, impeding the activation of the 'blue' genes at the appropriate stage, similar to the case of *Evx2* in the normal situation¹⁸. In such a case, *Shh* transcripts would be absent and the limb truncated. Therefore, the mechanism underlying *Hox* genes collinearity in early limb buds is critical for proper A–P limb polarity. Its co-option from trunk development makes the polarity of our appendages dependent on genetic constraints imposed by the general architecture of our body plan.

to these regulatory controls, determining its spatio-temporal expression. In many respects, this process resembles that implemented during the development of the major body axis, which also relies upon sequential activation of *Hox* genes in time and space²¹, as illustrated by common regulatory re-allocations in both early limb buds and extending tail buds observed in some cases after genetic alterations of the *HoxD* cluster²².

Consequently, we suggest that tetrapod limbs evolved along with the recruitment of the *Hox* collinear mechanism implemented in the developing body axis. This regulatory co-option imposed collinearity within limb buds, leading to the posterior-only expression of 5'-located *Hox* genes. Because these latter genes have the capacity to elicit *Shh* transcription, *Shh* signalling was confined to posterior limb bud cells, and hence the limb emerged with a built-in A–P polarity. Alternatively, all HOX proteins originally may have had the capacity to regulate *Shh*, this property being subsequently restricted to some genes only, to generate an A–P polarized structure. However, the absence of animals with appendages showing a truly bilateral symmetry, including ancestral tetrapod fossils^{19,23}, makes the latter scenario unlikely.

In the former view, the limb A–P polarity is a collateral effect of the genetic strategy co-opted for the distal extension of our limbs, a process highly constrained by the intrinsic logic of our body architecture. The necessity of restricting expression of 5'-located *Hox* genes caudally during gastrulation to prevent the deleterious effects of their products if expressed too rostrally in the main body axis (for example, see ref. 24), is thus probably the origin of their posterior expression in limb buds and of the consequent appendage A–P polarity. In this context, rather than the result of the selection of an independent regulatory strategy applied to an appendage because of the obvious adaptative advantages associated with such a polarity, we may consider that this polarity simply reflects the most parsimonious way of producing a limb.

METHODS

The production of the various mutant strains used in this study was previously described^{17,18,25}. Genotyping of mice and embryos was performed using Southern blot analysis. Whole-mount *in situ* hybridizations were carried out on 9.5- to 11-day-old fetuses, using standard procedures and the previously described *Shh* riboprobe²⁶. For skeletal preparations, adult or newborn animals were processed according to ref. 27. Embryos were carefully staged and the slight differences observed sometimes between the sizes of the limb buds reflect the phenotypic consequences of the genetic alterations rather than different ages.

Received 14 July; accepted 15 September 2006.

1. Saunders, J. W. & Gasselin, M. T. *Ectodermal-Mesodermal Interaction in the Origin of Limb Symmetry* 78–97 (Williams and Wilkins, Baltimore, 1968).
2. Riddle, R. D., Johnson, R. L., Laufer, E. & Tabin, C. *Sonic hedgehog* mediates the polarizing activity of the ZPA. *Cell* **75**, 1401–1416 (1993).
3. Knezevic, V. *et al.* *Hoxd-12* differentially affects preaxial and postaxial chondrogenic branches in the limb and regulates *Sonic hedgehog* in a positive feedback loop. *Development* **124**, 4523–4536 (1997).
4. Charite, J., de Graaff, W., Shen, S. & Deschamps, J. Ectopic expression of *Hoxb-8* causes duplication of the ZPA in the forelimb and homeotic transformation of axial structures. *Cell* **78**, 589–601 (1994).
5. Zakany, J., Kmita, M. & Duboule, D. A dual role for *Hox* genes in limb anterior-posterior asymmetry. *Science* **304**, 1669–1672 (2004).

6. Capellini, T. D. *et al.* *Pbx1/Pbx2* requirement for distal limb patterning is mediated by the hierarchical control of *Hox* gene spatial distribution and *Shh* expression. *Development* **133**, 2263–2273 (2006).
7. Laufer, E., Nelson, C. E., Johnson, R. L., Morgan, B. A. & Tabin, C. *Sonic hedgehog* and *Fgf-4* act through a signaling cascade and feedback loop to integrate growth and patterning of the developing limb bud. *Cell* **79**, 993–1003 (1994).
8. Niswander, L., Jeffrey, S., Martin, G. R. & Tickle, C. A positive feedback loop coordinates growth and patterning in the vertebrate limb. *Nature* **371**, 609–612 (1994).
9. Zuniga, A., Haramis, A. P., McMahon, A. P. & Zeller, R. Signal relay by BMP antagonism controls the SHH/FGF4 feedback loop in vertebrate limb buds. *Nature* **401**, 598–602 (1999).
10. Sun, X. *et al.* Conditional inactivation of *Fgf4* reveals complexity of signalling during limb bud development. *Nature Genet.* **25**, 83–86 (2000).
11. Khokha, M. K., Hsu, D., Brunet, L. J., Dionne, M. S. & Harland, R. M. Gremlin is the BMP antagonist required for maintenance of *Shh* and *Fgf* signals during limb patterning. *Nature Genet.* **34**, 303–307 (2003).
12. Scherz, P. J., Harfe, B. D., McMahon, A. P. & Tabin, C. J. The limb bud *Shh-Fgf* feedback loop is terminated by expansion of former ZPA cells. *Science* **305**, 396–399 (2004).
13. Chiang, C. *et al.* Manifestation of the limb prepattern: limb development in the absence of *Sonic hedgehog* function. *Dev. Biol.* **236**, 421–435 (2001).
14. Kraus, P., Fraidenreich, D. & Loomis, C. A. Some distal limb structures develop in mice lacking *Sonic hedgehog* signaling. *Mech. Dev.* **100**, 45–58 (2001).
15. Lewis, P. M. *et al.* Cholesterol modification of *Sonic hedgehog* is required for long-range signaling activity and effective modulation of signaling by *Ptc1*. *Cell* **105**, 599–612 (2001).
16. Harfe, B. D. *et al.* Evidence for an expansion-based temporal *Shh* gradient in specifying vertebrate digit identities. *Cell* **118**, 517–528 (2004).
17. Kmita, M. *et al.* Early developmental arrest of mammalian limbs lacking *HoxA/HoxD* gene function. *Nature* **435**, 1113–1116 (2005).
18. Tarchini, B. & Duboule, D. Control of *Hoxd* genes' collinearity during early limb development. *Dev. Cell* **10**, 93–103 (2006).
19. Shubin, N. H., Tabin, C. J. & Carroll, S. Fossils, genes and the evolution of animal limbs. *Nature* **388**, 639–648 (1997).
20. Fromental-Ramain, C. *et al.* Specific and redundant functions of the paralogous *Hoxa-9* and *Hoxd-9* genes in forelimb and axial skeleton patterning. *Development* **122**, 461–472 (1996).
21. Izpisua-Belmonte, J. C. *et al.* Murine genes related to the *Drosophila* *AbdB* homeotic gene are sequentially expressed during development of the posterior part of the body. *EMBO J.* **10**, 2279–2289 (1991).
22. van der Hoeven, F., Zakany, J. & Duboule, D. Gene transpositions in the *HoxD* complex reveal a hierarchy of regulatory controls. *Cell* **85**, 1025–1035 (1996).
23. Coates, M. I. The origin of vertebrate limbs. *Development (Suppl.)* **120**, 169–180 (1994).
24. Kmita, M., van der Hoeven, F., Zakany, J., Krumlauf, R. & Duboule, D. Mechanisms of *Hox* gene collinearity: transposition of the anterior *Hoxb1* gene into the posterior *HoxD* complex. *Genes Dev.* **14**, 198–211 (2000).
25. Tarchini, B., Huynh, T. H., Cox, G. A. & Duboule, D. *HoxD* cluster scanning deletions identify multiple defects leading to paralysis in the mouse mutant *Ironside*. *Genes Dev.* **19**, 2862–2876 (2005).
26. Echelard, Y. *et al.* *Sonic hedgehog*, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity. *Cell* **75**, 1417–1430 (1993).
27. Inouye, M. Differential staining of cartilage and bone in fetal mouse skeleton by alcian blue and alizarin red. *Congenital Anomalies* **16**, 171–173 (1976).

Acknowledgements We thank N. Fraudeau and T. H. N. Huynh for technical assistance as well as C. Tabin and M. Coates for comments and suggestions. This work was supported by funds from the canton de Genève, the Louis-Jeantet foundation, the Swiss National Research Fund, the National Center for Competence in Research (NCCR) 'Frontiers in Genetics' and the EU programme 'Cells into Organs'.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to D.D. (denis.duboule@zoo.unige.ch).

Effects of biodiversity on the functioning of trophic groups and ecosystems

Bradley J. Cardinale¹, Diane S. Srivastava², J. Emmett Duffy³, Justin P. Wright⁴, Amy L. Downing⁵, Mahesh Sankaran⁶ & Claire Jouseau⁷

Over the past decade, accelerating rates of species extinction have prompted an increasing number of studies to reduce species diversity experimentally and examine how this alters the efficiency by which communities capture resources and convert those into biomass^{1,2}. So far, the generality of patterns and processes observed in individual studies have been the subjects of considerable debate^{3–7}. Here we present a formal meta-analysis of studies that have experimentally manipulated species diversity to examine how it affects the functioning of numerous trophic groups in multiple types of ecosystem. We show that the average effect of decreasing species richness is to decrease the abundance or biomass of the focal trophic group, leading to less complete depletion of resources used by that group. At the same time, analyses reveal that the standing stock of, and resource depletion by, the most species-rich polyculture tends to be no different from that of the single most productive species used in an experiment. Of the known mechanisms that might explain these trends, results are most consistent with what is called the ‘sampling effect’, which occurs when diverse communities are more likely to contain and become dominated by the most productive species. Whether this mechanism is widespread in natural communities is currently controversial. Patterns we report are remarkably consistent for four different trophic groups (producers, herbivores, detritivores and predators) and two major ecosystem types (aquatic and terrestrial). Collectively, our analyses suggest that the average species loss does indeed affect the functioning of a wide variety of organisms and ecosystems, but the magnitude of these effects is ultimately determined by the identity of species that are going extinct.

Whereas one of the most striking features of our planet is its great variety of life, one of the most pervasive environmental changes of our time is the global loss of this biological diversity^{8,9}. Considerable uncertainty exists about current rates of extinction, but estimates place it somewhere between two and three orders of magnitude higher than rates found in the fossil record^{10,11}. Biologists have long pondered the environmental effects of species extinction. Even so, it was not until the 1990s that research efforts began to formalize the hypothesis that species diversity might influence the fluxes of energy and matter that are fundamental to all ecological processes, including those that control the abundance, biomass and distribution of organisms. Seminal studies suggested that species loss does, in fact, decrease how productive communities are and how efficiently they capture and consume limited resources^{12–14}. But the interpretation of these studies provoked considerable debate^{3–7}, and subsequent work produced several counterexamples that questioned the generality of these biodiversity effects^{15–19}. As a result, it has been argued that the

consequences of biodiversity loss are likely to be idiosyncratic, differing quantitatively and qualitatively between trophic groups and ecosystems^{20–23}.

After more than a decade of research, a sufficient number of studies have now emerged to permit rigorous testing of whether there are indeed general effects of biodiversity on ecosystem functioning. Here we present a formal meta-analysis of 111 field, greenhouse and laboratory experiments that have manipulated the diversity of species for a wide variety of organisms and ecosystems (see Supplementary Information). We focused on experiments that varied the richness of three or more species in a given trophic group t and measured either of two response variables: the aggregate abundance or biomass of all species in t (referred to as ‘standing stock’) and/or the total amount of resources depleted by t from a known resource pool (see Methods). Data were summarized for four trophic groups: first, microalgal, macroalgal or herbaceous plants assimilating nutrients or water; second, protozoan or metazoan herbivores consuming live algal or herbaceous plant tissue; third, protozoan or metazoan predators consuming live prey; and fourth, bacterial, fungal or metazoan detritivores consuming dead organic matter. Diversity effects were quantified with two complementary metrics. First, for each experiment i , we calculated the proportional difference in the response variable y between the mean value of the most species-rich polyculture $\bar{y}$ and the mean value of these same species grown in monoculture $\bar{m}$ as the log response ratio $LR_{\bar{m}} = \ln(y_{\bar{y}}/y_{\bar{m}})$. This unitless metric allows us to test whether there is a significant change in y with increasing richness when averaged across all species used in an experiment. We then calculated a complementary metric that quantifies the proportional difference between the mean value of the most species-rich polyculture and that of the taxon having the highest (lowest) mean value of y in monoculture $\bar{m}$, as $LR_{\bar{m}} = \ln(y_{\bar{y}}/y_{\bar{m}})$, where $y_{\bar{m}}$ is the highest (lowest) value when $LR_{\bar{m}} > 0$ (< 0). Testing whether $LR_{\bar{m}} > 0$ is analogous to tests for ‘transgressive’ overyielding, which are widely used to assess whether diverse polycultures are any more productive than the single most productive species²⁴.

Our analyses reveal quite general and consistent mean effects of species diversity on the aggregate abundance or biomass of species in a trophic group, with cascading effects on the resources used by that group. For $LR_{\bar{m}}$, we found that species richness positively affected the standing stocks of all four trophic groups considered, increasing the abundance or biomass of plants, herbivores, predators and detritivores (Fig. 1a). Higher diversity within each group was also associated with more complete depletion of resources (Fig. 1b). Experimentally increasing plant, predator and detritivore diversity all led to greater decreases in nutrients/water, prey, and dead organic

¹Department of Ecology, Evolution and Marine Biology, University of California at Santa Barbara, Santa Barbara, California 93106, USA. ²Department of Zoology, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada. ³Virginia Institute of Marine Science, The College of William and Mary, Gloucester Point, Virginia 23062, USA. ⁴Department of Biology, Duke University, Durham, North Carolina 27708, USA. ⁵Department of Zoology, Ohio Wesleyan University, Delaware, Ohio 43015, USA. ⁶Institute of Integrative & Comparative Biology, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK. ⁷Department of Ecology, Evolution and Environmental Biology, Columbia University, New York, New York 10027, USA.

matter, respectively. There was a similar tendency for increasing herbivore diversity to decrease the amount of living plant matter ($P = 0.08$ for a mixed-model analysis of variance; see Methods). In total, 67 of 76 experiments recorded positive values of LR_m for standing stock (88%), and 54 of 70 recorded positive values for resource depletion (77%).

When LR_m was modelled as a function of trophic group, we found no significant difference in the average diversity effect size between the four trophic groups for either response variable (Table 1). Furthermore, we found no significant difference in the average diversity effect size between studies performed in aquatic and terrestrial ecosystems (Fig. 1c, d and Table 1). This degree of consistency is remarkable given that the experiments spanned a wide variety of life forms (bacteria, fungi, plants and animals) and many of Earth's major ecosystems (lakes, streams, oceanic coastal habitat, temperate grasslands and forests; see Supplementary Information). Although studies are certainly not invariable in their conclusions, our results suggest that variation among studies is not consistent with previously proposed differences between trophic levels or ecosystems^{19–22}.

One of the major controversies in biodiversity research concerns the fact that some species exert stronger control over ecological processes than others³. Thus, a primary question when interpreting the average effect of species diversity is whether a diverse polyculture performs any differently than the single 'best' species (that is, the species having the greatest influence over a process). Our analyses show that the standing stock of, and resource depletion by, the most diverse species polyculture is statistically indistinguishable from that of the single species that achieves the highest level of these response variables in monoculture. Specifically, LR_m did not differ from zero for any of the four trophic groups (Fig. 1a, b) or for either of the two ecosystems (Fig. 1c, d). These conclusions hold true even if we apply a liberal test, considering only studies in which $LR_m > 0$ ($P = 0.13$ for standing stock, 0.27 for resource depletion). Of the known mechanisms by which species diversity can affect ecosystem functioning, these results are most consistent with what is called the 'sampling effect' of biodiversity, in which communities comprising more species have a greater chance of being dominated by the most productive

taxa. Note, however, that confirmation of this mechanism requires data on the covariance between competitive dominance in polyculture and the performance of species in monoculture^{24,25}—data that are not generally reported. There has been much controversy about whether the sampling effect is best interpreted as a 'real' biological mechanism that operates in nature or as an artefact of experiments that use random draws of species to assemble experimental communities^{1,3,6}. Until this debate is resolved, the relevance of the sampling effect for predicting the functional consequences of extinction is open to debate.

Our use of log response ratios to quantify the effects of species diversity could be criticized on grounds that these ratios compare only two ends of a continuum (highest versus lowest diversity). Because the highest levels of diversity differ between experiments (range 3–72 species) and tend to be higher in studies of terrestrial organisms than in those of aquatic organisms ($t = 4.64$, $P < 0.01$; 12.2 ± 9.6 species for terrestrial studies, 5.4 ± 8.1 for aquatic studies (means $\pm$ s.d.)), it is useful to ask how the general form of the diversity effect changes across levels of species richness. For 57 of 76 experiments that measured standing stock of a trophic group, and 51 of 70 experiments that measured resource depletion, species were manipulated at three or more levels of richness. This allowed us to fit data from each study to the Michaelis–Menten function $Y = Y_{\max}S/(K + S)$, where Y is the standing stock of, or resource depletion by, a trophic group standardized relative to the mean value of all monocultures y_m (that is, $Y = y_s/y_m$ where y_s is the value of y at richness level S). $Y_{\max}$ is therefore the maximum proportion by which Y increases or decreases relative to the average one-species system, and K describes how quickly Y approaches $Y_{\max}$ with increasing diversity. This function was an excellent fit to the data (median $R^2 = 0.84$), and better than several other models (see Methods). Thus, we used maximum-likelihood estimates of $Y_{\max}$ and K to compare key features of the diversity–function curves across systems.

We found no significant differences in $Y_{\max}$ or K between aquatic and terrestrial ecosystems (Fig. 2 and Table 1), which indicates that

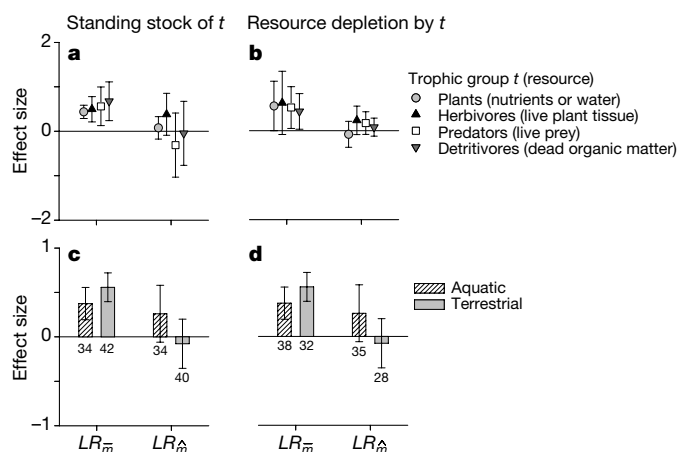


Figure 1 | Effects of species richness on the standing stock abundance or biomass of trophic group t , and the depletion of resources consumed by t . Data are means and 95% CI for two log response ratios that estimate the diversity 'effect size' from experiments. LR_m compares the mean value of the response variable y in a polyculture with the mean value of y averaged across the same species in monoculture. $LR_m > 0$ indicates that more diverse polycultures achieve higher standing stock (a, c) and deplete resources more fully (b, d) than the average monoculture. LR_h compares the mean value of y in a polyculture with that of the species with the highest (for $LR_m > 0$) or lowest (for $LR_m < 0$) mean value of y in monoculture (see the text). $LR_m = 0$ indicates that polycultures perform no differently than monocultures of the most productive species. Results are divided between four trophic groups (a, b) and two ecosystem types (c, d).

Table 1 | Statistical comparison of diversity effect sizes

Variable	Among trophic groups		
	d.f.	F	Pr > F
LR_m			
Standing stock of t	76	0.43	0.73
Resource depletion by t	70	0.09	0.96
LR_h			
Standing stock of t	74	0.98	0.41
Resource depletion by t	63	0.86	0.47
Variable	Among ecosystems		
	d.f.	F	Pr > F
LR_m			
Standing stock of t	76	2.27	0.14
Resource depletion by t	70	0.13	0.72
LR_h			
Standing stock of t	74	3.24	0.08
Resource depletion by t	63	1.62	0.21
Standing stock of t			
$Y_{\max}$	55	1.93	0.17
K	55	2.74	0.10
Resource depletion by t			
$Y_{\max}$	50	1.73	0.19
K	50	1.09	0.30

Results for LR_m and LR_h are from separate mixed-model analyses of variance that compare how richness in trophic group t influences the standing stock abundance or biomass of t , and the depletion of resources consumed by t among trophic groups (plants, herbivores, predators and detritivores) or ecosystems (aquatic and terrestrial). LR_m is the log ratio comparing the mean value of the response variable y in a polyculture with the mean value of y for the same species in monoculture. LR_h is the log ratio comparing the mean value of y in a polyculture with that of the species having the highest (for $LR_m > 0$) or lowest (for $LR_m < 0$) mean value of y in monoculture (see the text). Results for $Y_{\max}$ and K compare the maximum-likelihood parameter estimates for curves characterizing the diversity–function relationship (see Fig. 2) in aquatic and terrestrial ecosystems. All F values are non-significant, indicating that the effects of species richness on standing stocks and resource depletion do not differ between trophic groups or ecosystems.

the qualitative form of the diversity–function relationship is consistent across these habitat types (data were insufficient to make similar comparisons among trophic groups). With few exceptions, the curves were positive but decelerating, with values of $Y_{\max}$ being greater than the null expectation of unity (that is, $y_S > y_m$, $t = 14.0$ and $t = 11.9$ for standing stock and resource depletion, respectively, both $P < 0.01$) and values of K being greater than 0 ($t = 6.7$ and $t = 5.5$ for standing stock and resource depletion, respectively, both $P < 0.01$). Asymptotic estimates of $Y_{\max}$ suggest that the most diverse species polyculture would achieve 1.9-fold the standing stock of the average monoculture (95% confidence interval (CI) 1.6–2.2) and 1.8-fold the resource depletion (95% CI 1.5–2.1). Estimates of K indicate that half of the maximum value for both standing stock and resource depletion is achieved by the average species monoculture (mean for standing stock, 0.98 (95% CI 0.68–1.28); mean for resource depletion, 0.89 (95% CI 0.56–1.22)). However, the decelerating nature of these curves suggest that although a small number of species can maintain more than half the function, a disproportionately high number of species is required to maintain functions near maximal values.

Thus, our meta-analysis of 111 experiments conducted over more than a decade reveals two consistent results. First, as researchers have experimentally reduced the richness of species of a variety of organisms inhabiting numerous types of ecosystems, the average effect of diversity loss is to decrease the abundance or biomass of the focal trophic group, leading to less complete depletion of resources used by that group. Second, it is equally general that these average effects of species diversity on ecosystem functioning are best explained by the loss of the most productive species from a diverse community. There are at least two implications of these findings. First, from the perspective of basic research, our results present a new challenge to biologists. A fundamental tenet of biodiversity theory is that species must use resources in different ways to coexist stably^{26,27}. When

species do coexist by such niche differentiation, theory predicts that diverse polycultures will produce more biomass and capture a greater fraction of limited resources than even the ‘best’ species monoculture^{28,29}. The balance of evidence from experiments does not seem to support this, and understanding why there is a divergence between empirical and theoretical conclusions is one of the foremost challenges in this field. It may be that experiments have been performed at smaller spatial or shorter temporal scales than are the focus of theory, or that experiments do not meet equilibrium assumptions of theory. Second, our results re-emphasize a long-standing dilemma in the field of conservation biology—one that must soon be resolved. Biologists have long known that certain species exert much stronger control over ecological processes than others, but predicting which species these are in advance of extinction has proven difficult at best. A key challenge for future research is to detail more accurately how the traits that determine vulnerability to extinction are related to functional dominance in communities. Until that time, our finding that key aspects of ecosystem functioning decline consistently with the average species loss suggests that a precautionary approach to preserving as much biodiversity as possible is warranted.

METHODS

Selection of studies. We searched the literature for studies that experimentally manipulated the richness of three or more species in a given trophic group t and then measured a direct effect of richness on the standing stock of all species in t and/or the total depletion of resources by t . Standing stocks were calculated as the aggregate abundance or biomass of all organisms in t per unit area or volume. Depletion of resources was calculated as an instantaneous rate of consumption (for example, metabolic estimates of consumption of organic matter by bacteria or fungi), the difference between a known initial and a measured final resource concentration (for example, the depletion of soil nitrogen by plants), or the difference between treatments and zero-species controls (for example, the capture of prey by predators). Because our focus was on how richness affects ecosystem functioning at a given moment, we did not include studies focusing on community stability (that is, how diversity affects temporal variation in a dependent variable or invasibility). In all, we reviewed 184 papers, amassing data from digitized figures or tables or by acquiring original data from the authors of 58 studies reporting results from 111 experiments that met our criteria (see Supplementary Information).

Analyses of the diversity ‘effect size’. We used two log response ratios to quantify the diversity effect size in each experiment (see the text for the equations). LR_m was used to characterize the mean effect of diversity, testing whether the average of all replicates from the highest diversity treatment was different from the average response of these same species when grown in monocultures. In contrast, $LR_{\bar{m}}$ was used to test whether the average response of the highest diversity treatment was any different from that of the species having the highest (if $LR_m > 0$) or lowest (if $LR_m < 0$) value in monoculture. Two or more replicates of each monoculture were run for 59 of 63 experiments that measured resource depletion (94%), and 62 of 74 experiments measuring standing stock (84%). For these, we used the average value of replicates in our calculation of LR_m . For the small remainder of studies that had only $N = 1$ replicate for each monoculture, we used the point estimate.

Log ratios are the most widely used metrics in meta-analyses for two reasons: first, they estimate a proportional difference between treatments that can be readily compared between studies, and second, they have sampling properties that are known to be normal and that are robust to bias from small sample sizes³⁰. Mixed-model analyses of variance were used to test whether log response ratios differed from zero and to compare the mean values of these response ratios between trophic groups and ecosystem types. The general statistical model was $y_i = \mu + \tau_i + b_i + \varepsilon_i$, where y_i is LR_m or $LR_{\bar{m}}$ for each response variable, τ_i is a fixed categorical effect (trophic group or ecosystem type), b_i is the random effect associated with experiment i (with errors that are distributed normally and independently, $N[0, \sigma_b^2]$), and ε_i is the residual error. An important decision in meta-analyses is whether to standardize effect sizes by the variance of an experiment, giving greater weight to studies with higher ‘certainty’. We performed analyses with and without weighting, and these led to identical conclusions. Here we present unweighted results because these allow the more realistic, but also more variable, field studies to have the same influence on our conclusions as greenhouse and laboratory studies that tend to have higher replication and smaller variance.

Curve fitting. To characterize the general form of diversity–function relationships, we fit data from each study to three nonlinear functions that have

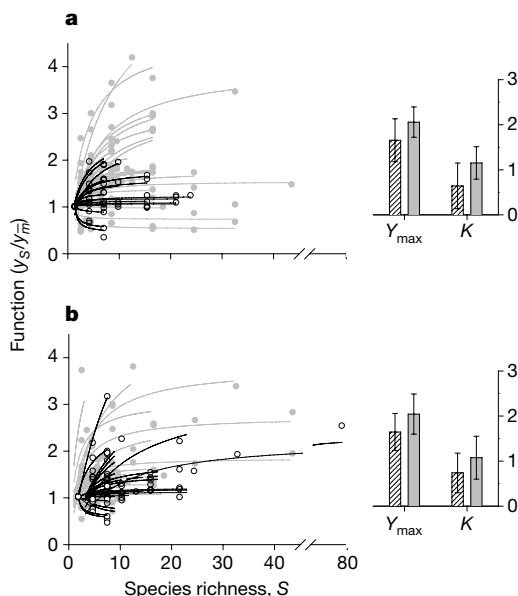


Figure 2 | The general form of the diversity–function relationship. Effects of species richness on the standing stock abundance or biomass of trophic group t (a), and the depletion of resources consumed by t (b). Each curve corresponds to data from a single study fitted to $Y = Y_{\max}S/(K + S)$, where Y is the proportional change in the dependent variable with increasing richness S , $Y_{\max}$ is the asymptotic estimate of Y , and K is the value of S at which $Y = Y_{\max}/2$. Sample sizes are 18 and 27 aquatic (black circles and lines), and 37 and 23 terrestrial studies (grey circles and lines) in a and b, respectively. Insets show the mean and 95% CI for the maximum-likelihood parameter estimates (hatched, aquatic; grey, terrestrial).

previously been used in the literature (log, power and hyperbolic). The Michaelis–Menten version of the hyperbolic function was the best-fitting model for the majority of studies (44% compared with 35% for power and 21% for log functions), and had the highest explanatory power (mean $R^2 = 0.71$, median 0.84). However, all three functions led to identical conclusions. For data fitted to the power function $\log(y) = m\log(S) + b$, m was positive (function increases with diversity, 95% CI 0.15–0.32 for standing stock and 0.11–0.29 for resource depletion) and did not differ between aquatic and terrestrial studies ($P > 0.26$ for both). For the log function $Y = b + m\log(S)$, m was positive (95% CI 0.25–0.66 for standing stock and 0.20–0.49 for resource depletion) and did not differ between aquatic and terrestrial studies ($P > 0.25$ for both).

Received 8 July; accepted 5 September 2006.

- Loreau, M. *et al.* Biodiversity and ecosystem functioning: Current knowledge and future challenges. *Science* **294**, 804–808 (2001).
- Chapin, F. S. *et al.* Biotic control over the functioning of ecosystems. *Science* **277**, 500–504 (1997).
- Huston, M. A. Hidden treatments in ecological experiments: Re-evaluating the ecosystem function of biodiversity. *Oecologia* **110**, 449–460 (1997).
- Schwartz, M. W. *et al.* Linking biodiversity to ecosystem function: Implications for conservation ecology. *Oecologia* **122**, 297–305 (2000).
- Schlapfer, F. & Schmid, B. Ecosystem effects of biodiversity: A classification of hypotheses and exploration of empirical results. *Ecol. Appl.* **9**, 893–912 (1999).
- Hooper, D. U. *et al.* Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. *Ecol. Monogr.* **75**, 3–35 (2005).
- Srivastava, D. S. & Vellend, M. Biodiversity–ecosystem function research: Is it relevant to conservation? *Annu. Rev. Ecol. Syst.* **36**, 267–294 (2006).
- Sala, O. E. *et al.* Global biodiversity scenarios for the year 2100. *Science* **287**, 1770–1774 (2000).
- Vitousek, P. M., Mooney, H. A., Lubchenco, J. & Melillo, J. M. Human domination of Earth's ecosystems. *Science* **277**, 494–499 (1997).
- Pimm, S. L., Russell, G. J., Gittleman, J. L. & Brooks, T. M. The future of biodiversity. *Science* **269**, 347–350 (1995).
- Millennium Ecosystem Assessment. *Ecosystems and Human Well-being: Biodiversity Synthesis* (World Resources Institute, Washington DC, 2005).
- Naeem, S., Thompson, L. J., Lawler, S. P., Lawton, J. H. & Woodfin, R. M. Declining biodiversity can alter the performance of ecosystems. *Nature* **368**, 734–737 (1994).
- Tilman, D., Wedin, D. & Knops, J. Productivity and sustainability influenced by biodiversity in grassland ecosystems. *Nature* **379**, 718–720 (1996).
- Hector, A. *et al.* Plant diversity and productivity experiments in European grasslands. *Science* **286**, 1123–1127 (1999).
- Wardle, D. A., Bonner, K. I. & Nicholson, K. S. Biodiversity and plant litter: Experimental evidence which does not support the view that enhanced species richness improves ecosystem function. *Oikos* **79**, 247–258 (1997).
- Downing, A. L. & Leibold, M. A. Ecosystem consequences of species richness and composition in pond food webs. *Nature* **416**, 837–841 (2002).
- Fridley, J. D. Resource availability dominates and alters the relationship between species diversity and ecosystem productivity in experimental plant communities. *Oecologia* **132**, 271–277 (2002).
- Finke, D. L. & Denno, R. F. Predator diversity dampens trophic cascades. *Nature* **429**, 407–410 (2004).
- Petchey, O. L., McPhearson, P. T., Casey, T. M. & Morin, P. J. Environmental warming alters food-web structure and ecosystem function. *Nature* **402**, 69–72 (1999).
- Raffaelli, D. *et al.* in *Biodiversity and Ecosystem Functioning: Synthesis and Perspectives* (eds Loreau, M., Naeem, S. & Inchausti, P.) 147–154 (Oxford Univ. Press, Oxford, 2002).
- Covich, A. P. *et al.* The role of biodiversity in the functioning of freshwater and marine benthic ecosystems. *Bioscience* **54**, 767–775 (2004).
- Emmerson, M. C., Solan, M., Emes, C., Paterson, D. M. & Raffaelli, D. Consistent patterns and the idiosyncratic effects of biodiversity in marine ecosystems. *Nature* **411**, 73–77 (2001).
- Duffy, J. E. Biodiversity loss, trophic skew and ecosystem functioning. *Ecol. Lett.* **6**, 680–687 (2003).
- Hector, A., Bazeley-White, E., Loreau, M., Otway, S. & Schmid, B. Overyielding in grassland communities: testing the sampling effect hypothesis with replicated biodiversity experiments. *Ecol. Lett.* **5**, 502–511 (2002).
- Loreau, M. & Hector, A. Partitioning selection and complementarity in biodiversity experiments. *Nature* **412**, 72–76 (2001).
- Hutchinson, G. E. Population studies—animal ecology and demography—concluding remarks. *Cold Spring Harb. Symp. Quant. Biol.* **22**, 415–427 (1957).
- Gause, G. F. *The Struggle for Existence* (Williams & Wilkins, Baltimore, Maryland, 1936).
- Tilman, D., Lehman, D. & Thompson, K. Plant diversity and ecosystem productivity: Theoretical considerations. *Proc. Natl Acad. Sci. USA* **94**, 1857–1861 (1997).
- Cardinale, B. J., Ives, A. R. & Inchausti, P. Effects of species diversity on the primary productivity of ecosystems: Extending our spatial and temporal scales of inference. *Oikos* **104**, 437–450 (2004).
- Hedges, L. V., Gurevitch, J. & Curtis, P. S. The meta-analysis of response ratios in experimental ecology. *Ecology* **80**, 1150–1156 (1999).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank S. Gaines, H. Hillebrand, M. Huston, J. Hille Ris-Lambers, J. Levine, J. Melack, B. Starzomski, D. Tilman and D. Wardle for comments that improved this manuscript. This work was supported by grants from the US National Science Foundation and is a product of the BioMERGE diversity-synthesis network.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to B.J.C. (cardinale@lifesci.ucsb.edu).

Corneal avascularity is due to soluble VEGF receptor-1

Balamurali K. Ambati^{1,2}, Miho Nozaki³, Nirbhai Singh¹, Atsunobu Takeda³, Pooja D. Jani¹, Tushar Suthar¹, Romulo J. C. Albuquerque³, Elizabeth Richter¹, Eiji Sakurai^{3,6}, Michael T. Newcomb³, Mark E. Kleinman³, Ruth B. Caldwell², Qing Lin⁷, Yuichiro Ogura⁶, Angela Orecchia⁸, Don A. Samuelson⁹, Dalen W. Agnew^{10,†}, Judy St. Leger¹¹, W. Richard Green¹², Parameshwar J. Mahasreshti¹³, David T. Curiel¹³, Donna Kwan¹⁴, Helene Marsh¹⁴, Sakae Ikeda¹⁵, Lucy J. Leiper¹⁶, J. Martin Collinson¹⁶, Sasha Bogdanovich¹⁷, Tejvir S. Khurana¹⁷, Masabumi Shibuya¹⁸, Megan E. Baldwin¹⁹, Napoleone Ferrara¹⁹, Hans-Peter Gerber^{19,†}, Sandro De Falco²⁰, Jassir Witta⁴, Judit Z. Baffi³, Brian J. Raisler^{3,5} & Jayakrishna Ambati^{3,5}

Corneal avascularity—the absence of blood vessels in the cornea—is required for optical clarity and optimal vision, and has led to the cornea being widely used for validating pro- and anti-angiogenic therapeutic strategies for many disorders^{1–4}. But the molecular underpinnings of the avascular phenotype have until now remained obscure^{5–10} and are all the more remarkable given the presence in the cornea of vascular endothelial growth factor (VEGF)-A, a potent stimulator of angiogenesis, and the proximity of the cornea to vascularized tissues. Here we show that the cornea expresses soluble VEGF receptor-1 (sVEGFR-1; also known as sflt-1) and that suppression of this endogenous VEGF-A trap¹¹ by neutralizing antibodies, RNA interference or Cre-lox-mediated gene disruption abolishes corneal avascularity in mice. The spontaneously vascularized corneas of *corn1* and *Pax6*^{+/-} mice^{12,13} and *Pax6*^{+/-} patients with aniridia¹⁴ are deficient in sflt-1, and recombinant sflt-1 administration restores corneal avascularity in *corn1* and *Pax6*^{+/-} mice. Manatees, the only known creatures uniformly to have vascularized corneas¹⁵, do not express sflt-1, whereas the avascular corneas of dugongs, also members of the order Sirenia, elephants, the closest extant terrestrial phylogenetic relatives of manatees, and other marine mammals (dolphins and whales) contain sflt-1, indicating that it has a crucial, evolutionarily conserved role. The recognition that sflt-1 is essential for preserving the avascular ambit of the cornea can rationally guide its use as a platform for angiogenic modulators, supports its use in treating neovascular diseases, and might provide insight into the immunological privilege of the cornea.

Although the absence of blood vessels in the cornea was known to the ancients, such as Susruta¹ and Galen², millennia ago, it was only in the last century that angiostatic substances were postulated to underpin corneal avascularity³. Because of its avascularity and ease of accessibility, the cornea has been a proving ground for anti-angiogenic strategies for more than 30 years⁴. However, the molecular foundations

of corneal avascularity remain unclear. In the last decade, many anti-angiogenic molecules such as angiostatin, endostatin, interleukin-1 receptor antagonist, pigment epithelium-derived factor, and thrombospondins have been found in the cornea (reviewed in ref. 5), leading to recognition of their actions as tumour suppressors, atherosclerotic plaque growth inhibitors, or modulators of wound healing. But none of these molecules is required for corneal avascularity, because mice deficient in any of them retain normal corneal phenotypes^{6–10}. This has led to a view of multiply redundant mechanisms of corneal avascularity.

The search for inhibitors of angiogenesis to treat atherosclerosis, cancer, diabetic kidney and retina damage, macular degeneration, and rheumatoid arthritis often relies on testing in the cornea for initial efficacy, owing to the absence of blood vessels in the cornea despite it being surrounded by the highly vascular conjunctiva (Fig. 1a). The cornea is ideal for understanding the ability of tissues to demarcate vascular ingrowth and identifying the efficacy of therapies against known angiogenic stimuli.

In our studies we found, surprisingly, that the cornea contained VEGF-A, but nearly all of it was bound (Fig. 1b). To understand the paradoxical presence of this potent pro-angiogenic molecule in an avascular tissue, we hypothesized that it was counterbalanced by the expression of sflt-1, an alternatively spliced, secreted isoform of the cell-surface receptor membrane-bound flt-1 (mbflt-1). sflt-1 lacks the transmembrane (tm) and tyrosine kinase (tk) domains of mbflt-1 and can act as a 'manacle' for VEGF-A (ref. 11). We found that sflt-1 mRNA and protein exist in the cornea (Fig. 1c–g, Supplementary Fig. 1); by contrast, mbflt-1 was found in the conjunctiva but not in the cornea (Fig. 1g). Consistent with its proposed function as a trap for secreted VEGF-A, sflt-1 was present extracellularly (Supplementary Fig. 2). We used immunoprecipitation to confirm that sflt-1 and VEGF-A interact *in vivo* (Fig. 1h) and corroborated it by immunostaining (Supplementary Fig. 3).

¹Departments of Ophthalmology and ²Cell Biology, Medical College of Georgia & Augusta Veterans Affairs Medical Center, Augusta, Georgia 30907, USA. ³Departments of Ophthalmology & Visual Sciences, ⁴Internal Medicine and ⁵Physiology, University of Kentucky, Lexington, Kentucky 40506, USA. ⁶Department of Ophthalmology, Nagoya City University Medical School, Nagoya 467-8601, Japan. ⁷Department of Microbiology and Immunology, Vanderbilt University School of Medicine, Nashville, Tennessee 37232, USA. ⁸Molecular and Cell Biology Laboratory, IDI-IRCCS, Rome 00167, Italy. ⁹Department of Small Animal Clinical Sciences, College of Veterinary Medicine, Gainesville, Florida 32610, USA. ¹⁰Department of Pathology, Microbiology and Immunology, University of California, Davis, California 95616, USA. ¹¹Department of Pathology, Sea World, San Diego, California 92109, USA. ¹²The Eye Pathology Laboratory, Wilmer Institute and Department of Pathology, Johns Hopkins Medical Institutions, Baltimore, Maryland 21205, USA. ¹³Division of Human Gene Therapy, The Gene Therapy Center, University of Alabama at Birmingham, Birmingham, Alabama 35294, USA. ¹⁴School of Tropical Environment Studies and Geography, James Cook University, Townsville, Queensland 4811, Australia. ¹⁵Department of Medical Genetics, University of Wisconsin, Madison, Wisconsin 53706, USA. ¹⁶School of Medical Sciences, University of Aberdeen, Foresterhill, Aberdeen AB25 2ZD, UK. ¹⁷Department of Physiology & Pennsylvania Muscle Institute, School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. ¹⁸Institute of Medical Science, University of Tokyo, Tokyo 108-8639, Japan. ¹⁹Department of Molecular Oncology, Genentech Inc., South San Francisco, California 94080, USA. ²⁰Institute of Genetics and Biophysics, Consiglio Nazionale delle Ricerche, Naples 80131, Italy. †Present addresses: Diagnostic Center for Population and Animal Health, Michigan State University, Lansing, Michigan 48910, USA (D.W.A.); Department of Translational Biology, Seattle Genetics, Bothell, Washington 98021, USA (H.-P.G.).

We used three independent strategies to test whether sflt-1 preserved corneal avascularity in mice. First, we injected a neutralizing antibody against flt-1 into the cornea, with fellow eyes receiving isotype control antibodies. Eyes that were treated with blocking antibody consistently developed corneal vascularization from the limbus within 1 day, whereas those treated with control antibody did not ($n = 14$, $P < 0.001$) (Supplementary Fig. 4a, b). Corneas treated with anti-flt-1 antibodies contained more free VEGF-A than did control-treated corneas (Supplementary Fig. 4c), indicating that sequestration of VEGF-A by sflt-1 maintains corneal avascularity. We confirmed this by showing that concomitantly treating corneas with neutralizing anti-VEGF-A antibodies, but not with isotype-control antibodies, prevented corneal vascularization induced by the anti-flt-1 antibody ($n = 8$, $P = 0.029$). Because the anti-flt-1 antibody would theoretically block ligand-binding by both mbflt-1 and sflt-1 (although the former is undetectable in the cornea), we tested this antibody in *flt-1* tyrosine kinase $^{-/-}$ (*flt-1* $tk^{-/-}$) mice, which are deficient in signalling induced by ligation of the flt-1 receptor¹⁶. The anti-flt-1 antibody, but not control antibodies, also induced

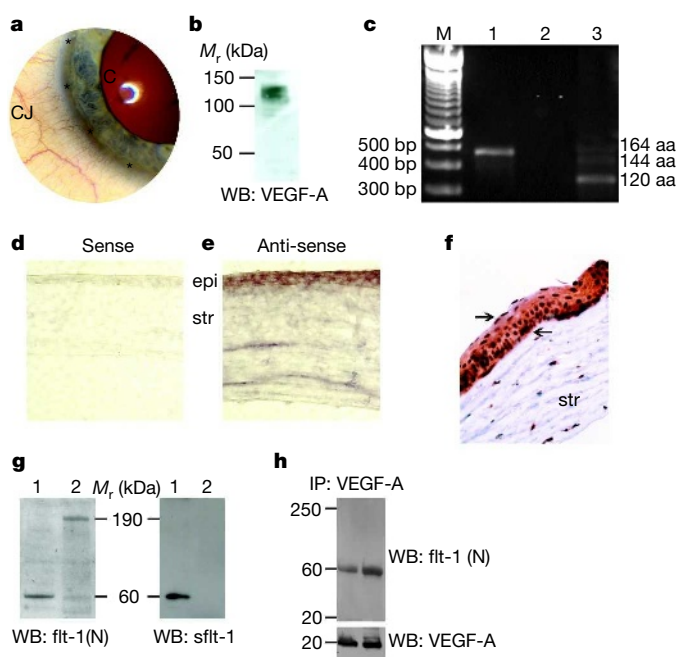


Figure 1 | Avascular cornea contains sflt-1 bound to VEGF-A. **a**, Photo of human eye demonstrates abrupt termination of blood vessels in the conjunctiva (CJ) at its border with the cornea (C), the limbus (asterisks). **b**, Representative non-reducing western blot of mouse cornea reveals immunoreactive bands of VEGF-A at 100–130 kDa corresponding to bound forms, and negligible immunoreactivity at 45–50 kDa corresponding to the free form. $n = 5$. **c**, sflt-1 (lane 1) and VEGF-A (lane 3) transcripts in mouse cornea identified by representative polymerase chain reaction with reverse transcription (RT-PCR). Lane 2 is water (template-negative) control. $n = 5$. bp, base pairs; aa, amino acids. **d**, **e**, sflt-1 mRNA detected by *in situ* hybridization in mouse corneal epithelium (epi) and stroma (str). Antisense RNA probes show purple-brown reactivity. Sense RNA probes show negligible reactivity. **f**, Immunolocalization (brown) of sflt-1 protein in mouse cornea. **g**, Representative reducing western blots (WB) with an antibody against the amino (N) terminus of flt-1 that recognizes both mbflt-1 and sflt-1 and an antibody against the unique C terminus of sflt-1 show that mouse cornea (1) contains primarily sflt-1 (60 kDa) whereas the conjunctiva (2) contains mainly mbflt-1 (190 kDa). **h**, Representative western blot of two independent mouse cornea samples immunoprecipitated (IP) with anti-VEGF-A antibody and immunoblotted (WB) with a biotinylated antibody against the N terminus of flt-1 that recognizes both mbflt-1 and sflt-1 shows that VEGF-A interacts with sflt-1 (60 kDa). Subsequent immunoblot with a biotinylated anti-VEGF-A antibody confirms the pull-down of VEGF-A by the immunoprecipitating antibody. $n = 6$.

corneal vascularization in *flt-1* $tk^{-/-}$ eyes ($n = 8$, $P = 0.029$), indicating that the vascular phenotype resulted from suppression of sflt-1 function and not from interference with flt-1 signalling. Subconjunctival injection of anti-flt-1 antibodies, which eliminates the confounding effect of corneal trauma, also elicited corneal vascularization ($n = 10$, $P = 0.008$; Supplementary Fig. 4d, e).

The second strategy was genomic deletion: we suppressed sflt-1 by conditional Cre-lox-mediated gene ablation because *flt-1* deletion is lethal¹⁷. Injections into the cornea of a plasmid¹⁸ that encoded Cre recombinase (pCre), but not of pNull, induced corneal vascularization in *flt-1*^{loxP/loxP} mice ($n = 10$; $P < 0.001$) within 2 days (Supplementary Fig. 5a, b). Expression of Cre was accompanied by significantly reduced sflt-1 and increased free VEGF-A (Supplementary Fig. 5c, d). Neither plasmid could induce corneal vascularization in wild-type mice ($n = 8$). To avoid injection trauma, we delivered a cell-permeable, enzymatically active Cre containing a nuclear localization sequence (NLS-Cre; refs 19, 20) to the cornea by topical eye drops (Fig. 2a–d). NLS-Cre, but not NLS-β-galactosidase, induced corneal vascularization in *flt-1*^{loxP/loxP} mice ($n = 11$; $P < 0.001$) within 2 days (Fig. 2e, f). Neither NLS-enzyme induced corneal vascularization in wild-type mice ($n = 8$; Fig. 2g, h).

The final strategy specifically knocked down sflt-1 using RNA interference (RNAi). We injected into the cornea a plasmid that expressed a short hairpin RNA (shRNA) targeted against a sequence in the unique carboxyl-terminus region of sflt-1 (pshRNA-sflt-1). The control was a plasmid that expressed an shRNA targeted against a sequence in the unique C-terminus region of mbflt-1 but that is not present in sflt-1 (pshRNA-mbflt-1). pshRNA-sflt-1, but not pshRNA-mbflt-1, substantially reduced sflt-1 mRNA and protein, indicating that knockdown was achieved through RNAi (Fig. 3a, b), and increased free VEGF-A (Fig. 3c), corroborating the thesis that sflt-1 sequesters VEGF-A to maintain physiological avascularity. Corneal vascularization was induced by pshRNA-sflt-1, but not pshRNA-mbflt-1, within 3 days of injection ($n = 36$, $P < 0.0001$) (Fig. 3d–f). pshRNA-sflt-1 also induced corneal vascularization in mice that had been systemically depleted of macrophages and neutrophils by treatment with clodronate liposomes and anti-Gr-1 antibodies (Supplementary Fig. 6a–c), indicating that corneal vascularization was not induced by infiltration of inflammatory cells and their delivery of VEGF-A. Furthermore, pshRNA-sflt-1 did not elevate VEGF-A mRNA (Supplementary Fig. 6d).

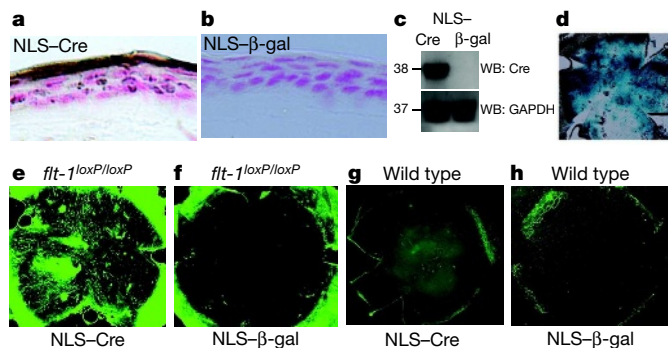


Figure 2 | Topical enzymatically active Cre recombinase abolishes corneal avascularity in *flt-1*^{loxP/loxP} mice. NLS-Cre but not NLS-β-galactosidase induces Cre expression (brown) in cornea within 1 h of eye drop application as shown by immunolocalization in cell nuclei stained red (**a**, **b**), and by representative reducing western blot (WB; **c**). **d**, X-gal staining of corneal flat mount of ROSA26R lacZ reporter mouse confirms expression of β-galactosidase (blue) 2 days after Cre expression. **e**, **f**, Representative corneal flat mounts show CD31⁺ (green)/LYVE-1[−] blood vessels in *flt-1*^{loxP/loxP} mouse corneas 14 days after treatment with NLS-Cre eye drops (**e**) but not with NLS-β-galactosidase (**f**). No corneal vascularization occurs in wild-type mice after topical NLS-Cre (**g**) or NLS-β-galactosidase (**h**). $n = 4$ (**a**, **b**, **d**), $n = 6$ (**c**), $n = 8–11$ (**e–h**).

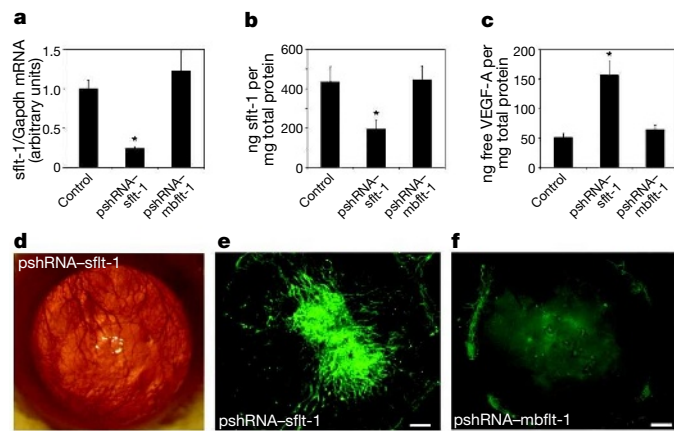


Figure 3 | Knockdown of sflt-1 mRNA abolishes corneal avascularity. **a–c**, Real-time RT–PCR reveals reduced sflt-1 mRNA (**a**) and enzyme-linked immunosorbent assay (ELISA) reveals reduced sflt-1 protein (**b**) and increased free VEGF-A protein (**c**) in wild-type mouse corneas 3 days after injection of pshRNA-sflt-1 but not pshRNA-mbflt-1. Asterisk denotes $P < 0.05$, Bonferroni corrected Mann–Whitney U -test. $n = 8–12$. Error bars depict s.e.m. **d–f**, pshRNA-sflt-1 but not pshRNA-mbflt-1 induces corneal vascularization in wild-type mice. $n = 36$. Photo of eye (**d**) and corneal flat mounts showing CD31⁺ (green, vascular endothelial cells) LYVE-1⁺ (lymphatic vessel endothelial hyaluronan receptor) blood vessels 14 days after injection (**e**, **f**). Scale bars, 500 μ m.

Like sflt-1, the transmembrane domain of flt-1 (flt-1-TM) also can trap VEGF-A, at least during development²¹. Like wild-type mice, *flt-1* $tk^{-/-}$ mice ($n > 60$), which retain expression of sflt-1 and flt-1-TM, have avascular corneas. Corneal vascularization was induced by pshRNA-sflt-1, but not pshRNA-mbflt-1, in *flt-1* $tk^{-/-}$ mice ($n = 8$; $P = 0.029$) just as in the wild type, indicating that sflt-1 and not flt-1-TM is required for corneal avascularity.

Apart from VEGF-A, sflt-1 also binds VEGF-B and placenta growth factor (PlGF). Expression of these ligands in mouse corneas was much less than that of VEGF-A (data not shown). Moreover, pshRNA-sflt-1, but not pshRNA-mbflt-1, induced corneal vascularization in both *Vegfb*^{-/-} ($n = 8$, $P = 0.029$) and *Plgf*^{-/-} ($n = 16$; $P < 0.0001$) mice, supporting the contention that corneal vascularization results from desequstration of VEGF-A from sflt-1. Direct evidence for this assertion was obtained by showing that corneal vascularization induced by pshRNA-sflt-1 in wild-type mice was prevented by a neutralizing anti-VEGF-A antibody but not by isotype-control antibodies ($n = 10$; $P = 0.008$).

Interferon (IFN)-mediated responses can allow pshRNAs to inhibit gene expression nonspecifically; however, pshRNA-sflt-1 induced corneal vascularization in mice deficient in IFN(alpha, beta and omega)receptor-1 (*Ifnar1*^{-/-}) or IFN(gamma) (*Ifng*^{-/-}) ($n = 8$) just as in wild-type mice, indicating that corneal vascularization was not attributable to IFN response effectors. To investigate whether other off-target effects might be responsible for corneal vascularization induced by pshRNA-sflt-1, we created p₂shRNA-sflt-1, which was targeted against a different sequence in the unique C-terminus region of sflt-1. Corneal injection of p₂shRNA-sflt-1 also induced corneal vascularization in wild-type mice ($n = 10$; Supplementary Fig. 7a), making it unlikely that off-target effects, which are sequence-specific and not target-specific, were responsible for loss of corneal avascularity.

To confirm that corneal vascularization induced by pshRNA-sflt-1 was mechanistically linked to sflt-1 knockdown, we developed a plasmid coding for a 'hardened-target' version of sflt-1 (psflt-1*). This contained seven translationally silent wobble position mutations that rendered expressed sflt-1 refractory to pshRNA-sflt-1. psflt-1*, but not psflt-1, prevented suppression of sflt-1 and the development of corneal vascularization in eyes that had been treated

with pshRNA-sflt-1 ($n = 10$, $P = 0.008$; Supplementary Fig. 7b). This functional control definitively established that the angiogenic phenotype was due to RNAi-mediated knockdown of sflt-1. Genetic, transcriptional and protein-targeting suppression of sflt-1 all induced corneal vascularization, showing that sflt-1 is the pre-eminent molecular defender of corneal avascularity.

The cornea remains avascular even in states of hypoxia such as those induced by eyelid closure during sleep or coma, and in a variety of ischaemic and occlusive disease states. We examined VEGF-A and sflt-1 levels in the corneas of mice exposed to 8% O₂ (comparable to corneal hypoxia during sleep) for 24 h. Despite profound hypoxia, these corneas remained avascular ($n = 20$). Although hypoxia can increase VEGF-A production, free VEGF-A was not significantly elevated in hypoxic corneas ($11 \pm 23\%$ greater than non-hypoxic corneas; $n = 9$; $P = 0.78$). This was attributed to an increase of $86 \pm 34\%$ in sflt-1 in hypoxic corneas ($n = 17$; $P = 0.05$), consistent with the presence of a hypoxia-responsive element in the flt-1 gene²². These data confer an important protective role upon sflt-1 in maintaining corneal avascularity during physiological hypoxia. By contrast, elevation of VEGF-A without concomitant induction of sflt-1, which was modelled by injection of recombinant VEGF-A, was reversed by administration of recombinant sflt-1/Fc but not of isotype control IgG/Fc (Supplementary Fig. 8), confirming its specificity.

We examined the spontaneously vascularized corneas of *corn1* and *Pax6*^{+/-} mice^{12,13} for the presence of sflt-1. Corneas of *corn1* and *Pax6*^{+/-} mice, unlike those of their background strains, were deficient in sflt-1 (Fig. 4a). Although it is unknown why these mice do not express sflt-1 in the cornea, it is notable that both strains have abnormalities in their corneal epithelium^{13,23}, the predominant source of sflt-1. sflt-1/Fc injection significantly reduced the area of corneal vascularization^{24,25} in *corn1* and *Pax6*^{+/-} mice compared with both IgG/Fc treated and untreated corneas (Fig. 4b, c), both implying that sflt-1 has a significant role in maintaining corneal homeostasis and suggesting that it might be possible to rescue corneal vascularization in a clinical setting. Although mutations in *destrin*, the protein that is altered in *corn1* mice, have not been reported in humans, *Pax6* mutations are present in patients with aniridia, who also have vascularized corneas¹⁴. Interestingly, the corneas of patients with aniridia ($n = 5$) were deficient in sflt-1 compared with normal human corneas ($n = 7$; Fig. 4d).

Florida manatees (*Trichechus manatus latirostris*) are the only organisms that have been reported uniformly to have spontaneously vascularized corneas¹⁵. We also observed this phenotype in the Antillean manatee (*Trichechus manatus manatus*; unpublished data). Interestingly, neither manatee expressed sflt-1 in the cornea, whereas the avascular corneas of dugongs (*Dugong dugon*), which also belong to the order Sirenia, and of Asian (*Elephas maximus*) and African (*Loxodonta africana*) elephants, the closest extant terrestrial phylogenetic relatives of manatees, did express sflt-1 in the cornea (Fig. 4e–g). The avascular corneas of other marine mammals such as dolphins (bottlenose: *Tursiops truncatus*; Risso's: *Grampus griseus*) and whales (Cuvier's beaked: *Ziphius cavirostris*; fin: *Balaenoptera physalus*; melon-headed: *Peponocephala electra*) also contained sflt-1 (Fig. 4h, i). The correlation between sflt-1 expression and corneal avascularity in diverse mammals supports the idea that sflt-1 has an evolutionarily conserved role in conferring corneal avascularity. Unlike dolphin and elephant corneas (Fig. 4i), manatee corneas expressed mbflt-1 (Supplementary Fig. 9), indicating that a splicing switch might account for their vascularized state. The teleological basis of the vascularized manatee cornea is intriguing. The absence of corneal sflt-1 and potentially suboptimal vision might result from a non-deleterious mutation in manatees, as they live primarily in turbid waters. Unlike dugongs, which are strictly marine, manatees are believed to be physiologically dependent on fresh water, and corneal vascularization could protect against, or perhaps result from, irritations due to this hypotonic environment.

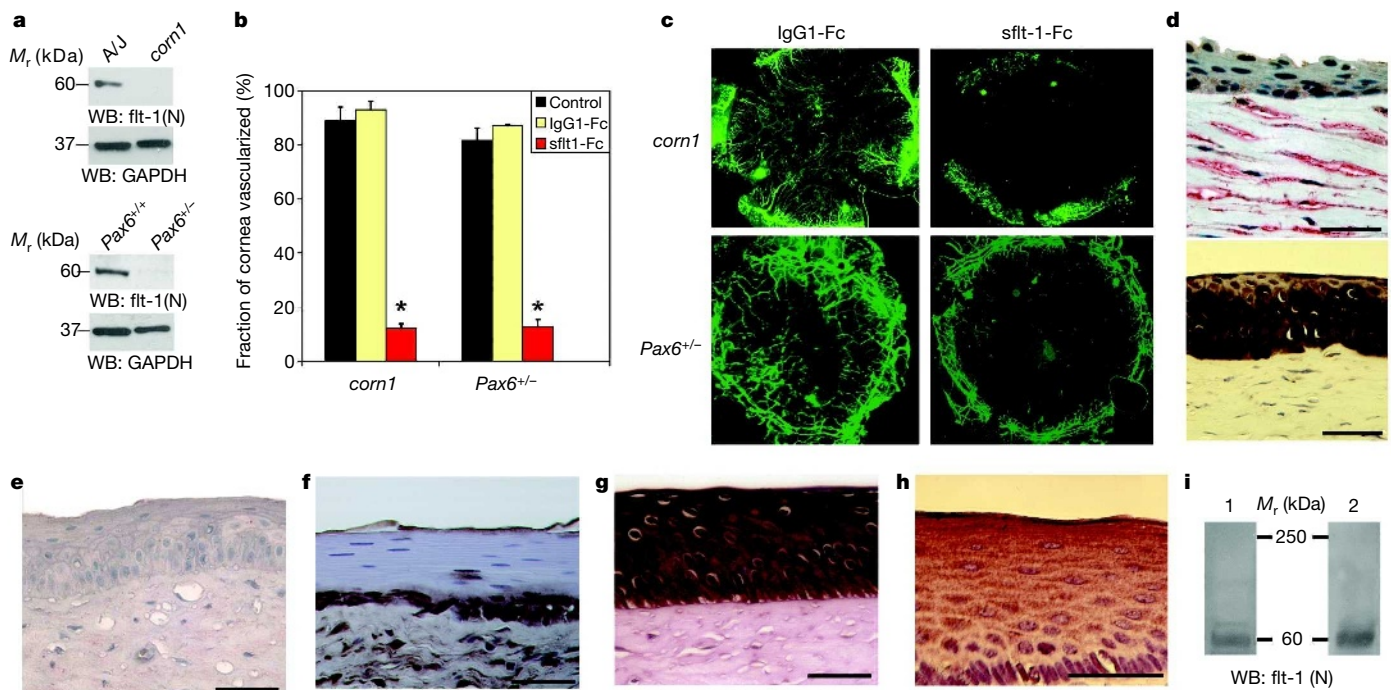


Figure 4 | Spontaneously vascularized corneas lacking sflt-1 are rescued by sflt-1 administration. **a**, Representative reducing western blots (WB) reveal deficiency of sflt-1 in corneas of *corn1* and *Pax6*^{+/-} mice compared with background strain A/J and *Pax6*^{+/+} mice. *n* = 10–12. **b**, **c**, Administration of sflt-1-Fc inhibits corneal vascularization in *corn1* and *Pax6*^{+/-} mice compared with administration of IgG1-Fc (by 87 ± 2% in *corn1*; *n* = 12; *P* = 0.01; by 85 ± 3% in *Pax6*^{+/-}; *n* = 10; *P* = 0.03) and with control untreated mice (by 87 ± 2% in *corn1*; *n* = 12; *P* = 0.01; by 84 ± 3% in *Pax6*^{+/-}; *n* = 10; *P* = 0.03). Significance by Bonferroni corrected Mann–Whitney *U*-test. Error bars depict s.e.m. (**b**). Representative flat mounts show CD31⁺ (green)/LYVE-1⁺ corneal blood vessels

The presence of many anti-angiogenic molecules in the cornea indicates that avascularity, which is essential for optical transparency and clear vision, might be maintained by multiply redundant mechanisms. Therefore the finding that neutralization or knock-down of sflt-1 alone abolishes corneal avascularity is surprising, but consistent with the presence of VEGF-A in the normal cornea. We speculate that VEGF-A is produced and held in a sequestered state by the cornea as a readily available store because this exposed tissue is susceptible to injuries that might require an angiogenic response. Alternatively, it might be a vestigial residue of an evolutionary requirement to provide blood to the eye that later required biochemical compensation in the form of sflt-1 expression to support improved vision.

The utilization of sflt-1 to regulate the bioavailability of VEGF-A is conserved in other systems such as cyclic vascularization²⁶ and embryonic sprouting²⁷, and disturbances in this regulation underlie preeclampsia²⁸. Our findings unveil a new role for sflt-1 in the evolutionary establishment of optimal vision resulting from and requiring optical clarity. Apart from trapping VEGF-A, sflt-1 can heterodimerize with mbflt-1 and VEGFR-2 (ref. 11). Although neither mbflt-1 nor VEGFR-2 is expressed in the normal cornea (Fig. 1g, Supplementary Fig. 10), such heterodimerization might modulate pathological vascularization of the cornea. Other mechanisms of regulating VEGF-A bioavailability, such as matrix metalloproteinase-induced release, have been identified in a model of tumour angiogenesis²⁹.

The cornea has long been used as a readout platform to assay anti-angiogenic therapy in oncology, cardiovascular biology and other fields. The recognition that sflt-1 is dominant in maintaining corneal avascularity directly affects the degree to which this tissue can be generalized in individual models. Our data also provide insights into

the relative immunological privilege of the cornea, as corneal avascularity is crucial to the high success rate of corneal allografts³⁰. These findings also support the use of sflt-1 in preventing or treating neovascularization. Furthermore, they illuminate its potential as a therapeutic target in conditions where inducing angiogenesis in an sflt-1-rich microenvironment might be beneficial, for example, preeclampsia, wound healing, stroke and heart disease.

METHODS

Imaging. *In vivo* images were captured by CCD camera (Nikon) under a dissecting microscope. Blood vessels were defined by positive labelling with FITC-conjugated rat antibody against mouse CD31 (1:333; BD Pharmingen) and negative labelling with rabbit antibody against mouse LYVE-1 (1:333; Abcam) on corneal flat mounts, as previously reported^{24,25}.

Injections. Neutralizing goat antibody (10 μg) against mouse flt-1 (R&D Systems), isotype control goat IgG (10 μg; Jackson ImmunoResearch), shRNAs (4 μg) against mbflt-1 or sflt-1, psflt-1 (4 μg), psflt-1* (4 μg), pCre (4 μg; gift of R.K. Nordeen, University of Colorado), pNull (4 μg), rmVEGF-A164 (20–500 ng; R&D Systems), sflt-1/Fc (5 μg; R&D Systems), or isotype control IgG1/Fc (5 μg; Jackson ImmunoResearch) were injected (2 μl) into the cornea with a 33-gauge needle, as previously reported¹⁸. The efficiency of corneal transfection by naked plasmid of pGFP (gift of X. Li, University of Kentucky) or placZ (gift of B.T. Spear, University of Kentucky) exceeded 70%, as gauged by flow cytometry and X-gal staining (Supplementary Fig. 11). We performed tail-vein injection of clodronate liposomes (200 μl) and intraperitoneal injection of anti-Gr-1 antibodies (200 μg; eBioscience) on each of the two days before and after corneal injection of pshRNA-sflt-1 to deplete peripheral monocytes/macrophages and neutrophils.

Received 19 June; accepted 15 September 2006.

Published online 18 October 2006.

- Sharma, P. V. *Susruta-Samhita* (Chaukhambha Visvabharati, Varanasi, India, 2001).

2. Magnus, H. *Ophthalmology of the Ancients* (J. P. Wayenborgh, Oostende, Belgium, 1999).
3. Meyer, K. & Chaffee, E. The mucopolysaccharide acid of the cornea and its enzymatic hydrolysis. *Am. J. Ophthalmol.* **23**, 1320–1325 (1940).
4. Gimbrone, M. A. Jr, Cotran, R. S., Leapman, S. B. & Folkman, J. Tumor growth and neovascularization: an experimental model using the rabbit cornea. *J. Natl Cancer Inst.* **52**, 413–427 (1974).
5. Chang, J. H., Gabison, E. E., Kato, T. & Azar, D. T. Corneal neovascularization. *Curr. Opin. Ophthalmol.* **12**, 242–249 (2001).
6. Wiegand, S. J. *et al.* Genetic modulation of pigment epithelium-derived factor (PEDF) expression does not alter normal or pathological angiogenesis in the eye, or tumor growth. *Invest. Ophthalmol. Vis. Sci.* **45**, abstr. 1884 (2004).
7. Cursiefen, C. *et al.* Roles of thrombospondin-1 and -2 in regulating corneal and iris angiogenesis. *Invest. Ophthalmol. Vis. Sci.* **45**, 1117–1124 (2004).
8. Bugge, T., Flick, M., Daugherty, C. & Degen, J. Plasminogen deficiency causes severe thrombosis but is compatible with development and reproduction. *Genes Dev.* **9**, 794–807 (1995).
9. Fukai, N. *et al.* Lack of collagen XVIII/endostatin results in eye abnormalities. *EMBO J.* **21**, 1535–1544 (2002).
10. Hirsch, E., Irikura, V. M., Paul, S. M. & Hirsh, D. Functions of interleukin 1 receptor antagonist in gene knockout and overproducing mice. *Proc. Natl Acad. Sci. USA* **93**, 11008–11013 (1996).
11. Kendall, R. L. & Thomas, K. A. Inhibition of vascular endothelial cell growth factor activity by an endogenously encoded soluble receptor. *Proc. Natl Acad. Sci. USA* **90**, 10705–10709 (1993).
12. Smith, R. *et al.* Corn1: a mouse model for corneal surface disease and neovascularization. *Invest. Ophthalmol. Vis. Sci.* **37**, 397–404 (1996).
13. Ramaesh, T. *et al.* Corneal abnormalities in Pax6^{+/−} small eye mice mimic human aniridia-related keratopathy. *Invest. Ophthalmol. Vis. Sci.* **44**, 1871–1878 (2003).
14. Jordan, T. *et al.* The human PAX6 gene is mutated in two patients with aniridia. *Nature Genet.* **1**, 328–332 (1992).
15. Harper, J. Y., Samuelson, D. A. & Reep, R. L. Corneal vascularization in the Florida manatee (*Trichechus manatus latirostris*) and three-dimensional reconstruction of vessels. *Vet. Ophthalmol.* **8**, 89–99 (2005).
16. Hiratsuka, S., Minowa, O., Kuno, J., Noda, T. & Shibuya, M. Flt-1 lacking the tyrosine kinase domain is sufficient for normal development and angiogenesis in mice. *Proc. Natl Acad. Sci. USA* **95**, 9349–9354 (1998).
17. Fong, G. H., Rossant, J., Gertsenstein, M. & Breitman, M. L. Role of the Flt-1 receptor tyrosine kinase in regulating the assembly of vascular endothelium. *Nature* **376**, 66–70 (1995).
18. Stechschulte, S. U. *et al.* Rapid ocular angiogenic control via naked DNA delivery to cornea. *Invest. Ophthalmol. Vis. Sci.* **42**, 1975–1979 (2001).
19. Lin, Q., Jo, D., Gebre-Amlak, K. D. & Ruley, H. E. Enhanced cell-permeant Cre protein for site-specific recombination in cultured cells. *BMC Biotechnol.* **4**, 25 (2004).
20. Jo, D. *et al.* Epigenetic regulation of gene structure and function with a cell-permeable Cre recombinase. *Nature Biotechnol.* **19**, 929–933 (2001).
21. Hiratsuka, S. *et al.* Membrane fixation of vascular endothelial growth factor receptor 1 ligand-binding domain is important for vasculogenesis and angiogenesis in mice. *Mol. Cell. Biol.* **25**, 346–354 (2005).
22. Gerber, H. P., Condorelli, F., Park, J. & Ferrara, N. Differential transcriptional regulation of the two vascular endothelial growth factor receptor genes. Flt-1, but not Flk-1/KDR, is up-regulated by hypoxia. *J. Biol. Chem.* **272**, 23659–23667 (1997).
23. Ikeda, S. *et al.* Aberrant actin cytoskeleton leads to accelerated proliferation of corneal epithelial cells in mice deficient for destrin (actin depolymerizing factor). *Hum. Mol. Genet.* **12**, 1029–1036 (2003).
24. Ambati, B. K. *et al.* Sustained inhibition of corneal neovascularization by genetic ablation of CCR5. *Invest. Ophthalmol. Vis. Sci.* **44**, 590–593 (2003).
25. Ambati, B. K., Jousen, A. M., Kuziel, W. A., Adamis, A. P. & Ambati, J. Inhibition of corneal neovascularization by genetic ablation of CCR2. *Cornea* **22**, 465–467 (2003).
26. Graubert, M. D., Ortega, M. A., Kessel, B., Mortola, J. F. & Iruela-Arispe, M. L. Vascular repair after menstruation involves regulation of vascular endothelial growth factor-receptor phosphorylation by sFLT-1. *Am. J. Pathol.* **158**, 1399–1410 (2001).
27. Kearney, J. B., Kappas, N. C., Ellerstrom, C., DiPaola, F. W. & Bautch, V. L. The VEGF receptor flt-1 (VEGFR-1) is a positive modulator of vascular sprout formation and branching morphogenesis. *Blood* **103**, 4527–4535 (2004).
28. Levine, R. J. *et al.* Circulating angiogenic factors and the risk of preeclampsia. *N. Engl. J. Med.* **350**, 672–683 (2004).
29. Bergers, G. *et al.* Matrix metalloproteinase-9 triggers the angiogenic switch during carcinogenesis. *Nature Cell Biol.* **2**, 737–744 (2000).
30. Dana, M. R. & Streilein, J. W. Loss and restoration of immune privilege in eyes with corneal neovascularization. *Invest. Ophthalmol. Vis. Sci.* **37**, 2485–2494 (1996).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the various aquaria, zoos and wildlife rehabilitation centres that donated tissues for comparative studies; R. Groom, S. Joshi, M. Kellogg, R. King, C. K. Lau, P. Lewis, N. Mezei, K.K. Smith and L. Xu for technical assistance; R. J. Kryscio for statistical guidance; and R. Mohan, S. Bondada, M. W. Fannon, L. Mazzaro, Y. Nozaki, P. A. Pearson, L. Peichl, A. M. Rao, G. S. Rao and K. Ambati for discussions. J.A. was supported by the NEI/NIH, the Lew R. Wasserman Merit Award (Research to Prevent Blindness), the Dennis W. Jahnigen Career Development Award (American Geriatrics Society, John A. Hartford Foundation, Atlantic Philanthropies), the Macula Vision Research Foundation, the International Retinal Research Foundation, the E. Matilda Ziegler Foundation for the Blind, the Dr E. Vernon Smith and Eloise C. Smith Macular Degeneration Endowed Chair, a physician–scientist award from University of Kentucky, and a departmental challenge grant from Research to Prevent Blindness; M.N. by ARVO/Japan National Society for the Prevention of Blindness; E.S. by Fight For Sight; R.J.C.A. by Research to Prevent Blindness; J.Z.B. and B.R. by the NIDCD/NIH; A.T. by a Japan Young Scientist Award; B.K.A. by a VA Career Development Award, the Knights-Templar Eye Foundation and Fight for Sight; N.S. by ARVO/Alcon; S.I. by the NEI/NIH; J.M.C. by the Wellcome Trust and the Birth Defects Foundation; T.S.K. by the NEI/NIH and the NIAMS/NIH; S.B. by the Muscular Dystrophy Association; and S.D.F. by the AIRC (Italian Association for Cancer Research).

Author Contributions B.K.A. and J.A. conceived and designed the experiments, wrote the manuscript, and are joint senior authors. M.N., N.S., A.T., P.D.J., J.Z.B. and B.J.R. contributed equally.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the paper on www.nature.com/nature. Correspondence and requests for materials should be addressed to B.K.A. (bambati@mail.mcg.edu) or J.A. (jamba2@uky.edu).

LETTERS

Chronic polyarthritis caused by mammalian DNA that escapes from degradation in macrophages

Kohki Kawane^{1,3,4}, Mayumi Ohtani^{1,4}, Keiko Miwa^{1,4†}, Takuji Kizawa², Yoshiyuki Kanbara⁵, Yoshichika Yoshioka⁵, Hideki Yoshikawa² & Shigekazu Nagata^{1,3,4}

A large amount of chromosomal DNA is degraded during programmed cell death and definitive erythropoiesis¹. DNase II is an enzyme that digests the chromosomal DNA of apoptotic cells and nuclei expelled from erythroid precursor cells after macrophages have engulfed them^{1,2}. Here we show that *DNase II*^{-/-} *IFN-IR*^{-/-} mice and mice with an induced deletion of the *DNase II* gene develop a chronic polyarthritis resembling human rheumatoid arthritis. A set of cytokine genes was strongly activated in the affected joints of these mice, and their serum contained high levels of anti-cyclic citrullinated peptide antibody, rheumatoid factor and matrix metalloproteinase-3. Early in the pathogenesis, expression of the gene encoding tumour necrosis factor (TNF)- α was upregulated in the bone marrow, and administration of anti-TNF- α antibody prevented the development of arthritis. These results indicate that if macrophages cannot degrade mammalian DNA from erythroid precursors and apoptotic cells, they produce TNF- α , which activates synovial cells to produce various cytokines, leading to the development of chronic polyarthritis.

DNase II^{-/-} mice die as embryos as a result of the constitutive production of interferon (IFN)- β ^{3,4}, and this lethality can be rescued by a deficiency of the *IFN-IR* gene⁵. *DNase II*^{-/-} *IFN-IR*^{-/-} mice were born normal, at a slightly reduced mendelian ratio, but they developed polyarthritis as they aged (see below). To examine whether mice carrying a single deletion in *DNase II* gene develop arthritis, we introduced a floxed allele into the *DNase II* gene (Supplementary Fig. S1). These mice were crossed with *Mx1-Cre* mice, which carry the *Cre* gene under the control of *Mx1*, an IFN-inducible promoter⁶, and *DNase II*^{lox/-} *Mx1-Cre*^T mice were established. When these mice were treated with poly(I)•poly(C) to induce IFN, the floxed region of *DNase II* gene was deleted within 2 months in the spleen (Supplementary Fig. S2), and *DNase II* messenger RNA was undetectable in the spleen and bone marrow (Supplementary Fig. S3). Hereafter we refer to the poly(I)•poly(C)-treated *DNase II*^{lox/-} mice as *DNase II*^{Δ/-} mice.

The inactivation of the *DNase II* gene in adult mice was not lethal. However, *DNase II*^{Δ/-}, but not poly(I)•poly(C)-treated

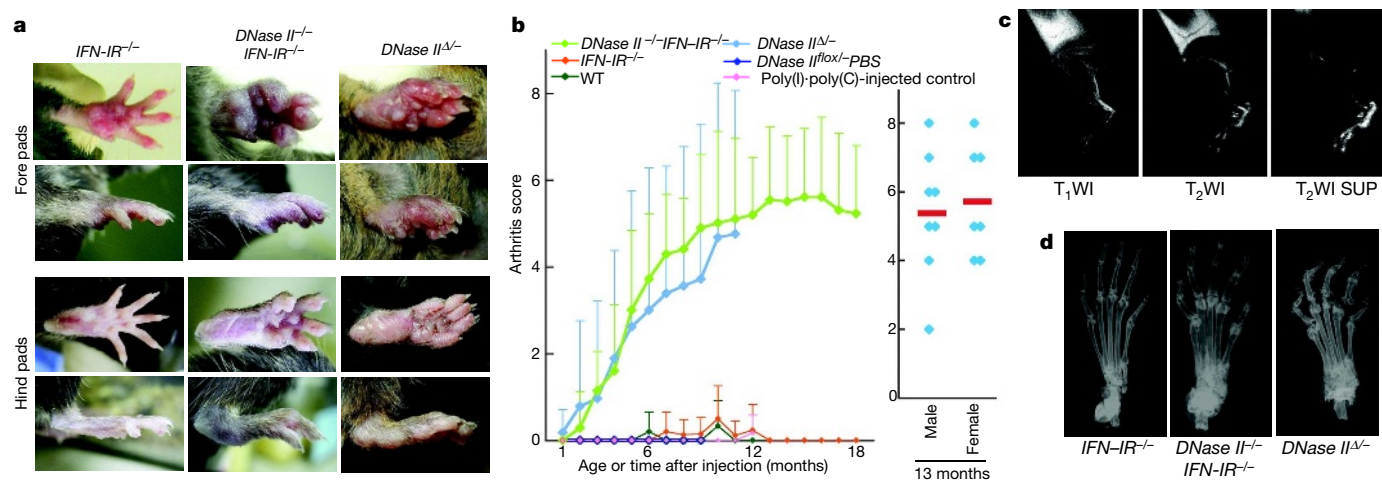


Figure 1 | Arthritis in *DNase II*^{-/-} *IFN-IR*^{-/-} and *DNase II*^{Δ/-} mice. **a**, Joint swelling of *DNase II*^{-/-} *IFN-IR*^{-/-} at the age of 12.5 months, and *DNase II*^{Δ/-} 8–10 months after treatment with poly(I)•poly(C). Joints of *IFN-IR*^{-/-} mice are shown as control. **b**, Left, joints of wild-type ($n = 3-4$), *IFN-IR*^{-/-} ($n = 5-8$) and *DNase II*^{-/-} *IFN-IR*^{-/-} ($n = 9-10$) mice were examined at the indicated ages. *DNase II*^{lox/-} ($n = 8-10$) and *DNase II*^{lox/+} ($n = 11-15$) littermates were given poly(I)•poly(C) or PBS ($n = 5$), and examined after the indicated periods. The arthritis scores are plotted as

means and s.d. Right, the scores of 13-month-old *DNase II*^{-/-} *IFN-IR*^{-/-} are plotted according to sex. The average values are indicated by horizontal red lines. **c**, MRI of the hindlimb of a 15-month-old *DNase II*^{-/-} *IFN-IR*^{-/-} mouse. T₁-weighted images (T₁WI), T₂-weighted images (T₂WI) and T₂-weighted images with fat suppression (T₂WISUP) are shown. **d**, Radiographic examination of the hindlimbs from a 15.5-month-old *DNase II*^{-/-} *IFN-IR*^{-/-} and a *DNase II*^{Δ/-} mouse 10 months after treatment with poly(I)•poly(C).

¹Department of Genetics and ²Department of Orthopaedics, Osaka University Medical School, Osaka 565-0871, Japan. ³Laboratory of Genetics, Integrated Biology Laboratories, Graduate School of Frontier Biosciences, Osaka University, Osaka 565-0871, Japan. ⁴Solution Oriented Research for Science and Technology, Japan Science and Technology Corporation, Osaka 565-0871, Japan. ⁵High Field Magnetic Resonance Imaging Research Institute, Advanced Medical Science Research Centre, Iwate Medical University, Takizawa 020-0173, Japan. [†]Present address: Laboratory of Cell Lineage Modulation, RIKEN Kobe Institute, Kobe 650-0047, Japan.

DNase II^{flox/+} mice, developed chronic polyarthritis in a time-dependent manner (Fig. 1). In *DNase II^{-/-}IFN-IR^{-/-}* mice at 2–3 months of age or in *DNase II^{flox/-}* mice 2–3 months after treatment with poly(I)•poly(C), forelimbs and hindlimbs began to swell. Swelling affected first the digit, then the foot, and finally the wrist and ankles. Individual mutant mice followed a significantly different time course, but all the mice were eventually affected. As swelling progressed, the mutant mice gradually lost their grasp strength, and their joints became deformed. There was no clear difference between male and female mice in disease development (Fig. 1b). Magnetic resonance imaging (MRI) of the affected joints showed high-intensity signals on the T₂-weighted image around tarsal bones and phalanges (Fig. 1c). These high-intensity regions were also observed on the fat-suppressed T₂-weighted image, indicating that the joints suffered massive inflammation. Radiographic examination revealed destruction, erosion, and deformity of the subchondral bones (Fig. 1d).

Histology of the swollen joints from both *DNase II^{-/-}IFN-IR^{-/-}* and *DNase II^{flox/-}* mice showed severe synovitis with villus proliferation accompanied by pannus formation (Fig. 2a). Pannus filled the joint cavity, eroded cartilage, destroyed bones, and occasionally penetrated the bone marrow. Pannus was dominated by macrophages carrying the CD68 antigen, among which tartrate-resistant acid phosphatase (TRAP)-positive osteoclasts were at the leading

edge. Infiltration of subsynovial tissues by T cells and neutrophils was also observed (Fig. 2a, and data not shown). The abnormal inflammation was limited to the joints, and no apparent disorder was found in the skin, liver, salivary glands, gut or blood vessels of most of the mice.

A set of cytokines and chemokines is known to be highly expressed in the joints of human patients with rheumatoid arthritis^{7–9}. Real-time polymerase chain reaction (PCR) analysis indicated that levels of mRNA for TNF- α , interleukin (IL)-1 β , IL-6, IL-10, IFN- β and IFN- γ in the affected joints of *DNase II^{-/-}IFN-IR^{-/-}* and *DNase II^{flox/-}* mice were 5–100-fold those in control mice (Fig. 2b). The level of mRNA for IL-18 in the joints of *DNase II^{-/-}* mice was comparable to that in control mice. However, the level of IL-18 protein was very high in the serum of *DNase II^{-/-}IFN-IR^{-/-}* and *DNase II^{flox/-}* mice (Fig. 2c), indicating that the expression of IL-18 might be regulated post-transcriptionally. Matrix metalloproteinase (MMP)-3 is produced by synovial cells of human patients with rheumatoid arthritis and seems to be involved in the degradation of the extracellular matrix¹⁰. Accordingly, the level of the mRNA for MMP-3 in the joints of *DNase II^{-/-}IFN-IR^{-/-}* and *DNase II^{flox/-}* mice was more than 25-fold that in control mice (Fig. 2b), and its protein level in the serum was also 2–3-fold higher in the mutant mice (Fig. 2c). Anti-cyclic citrullinated peptide (CCP) antibody,

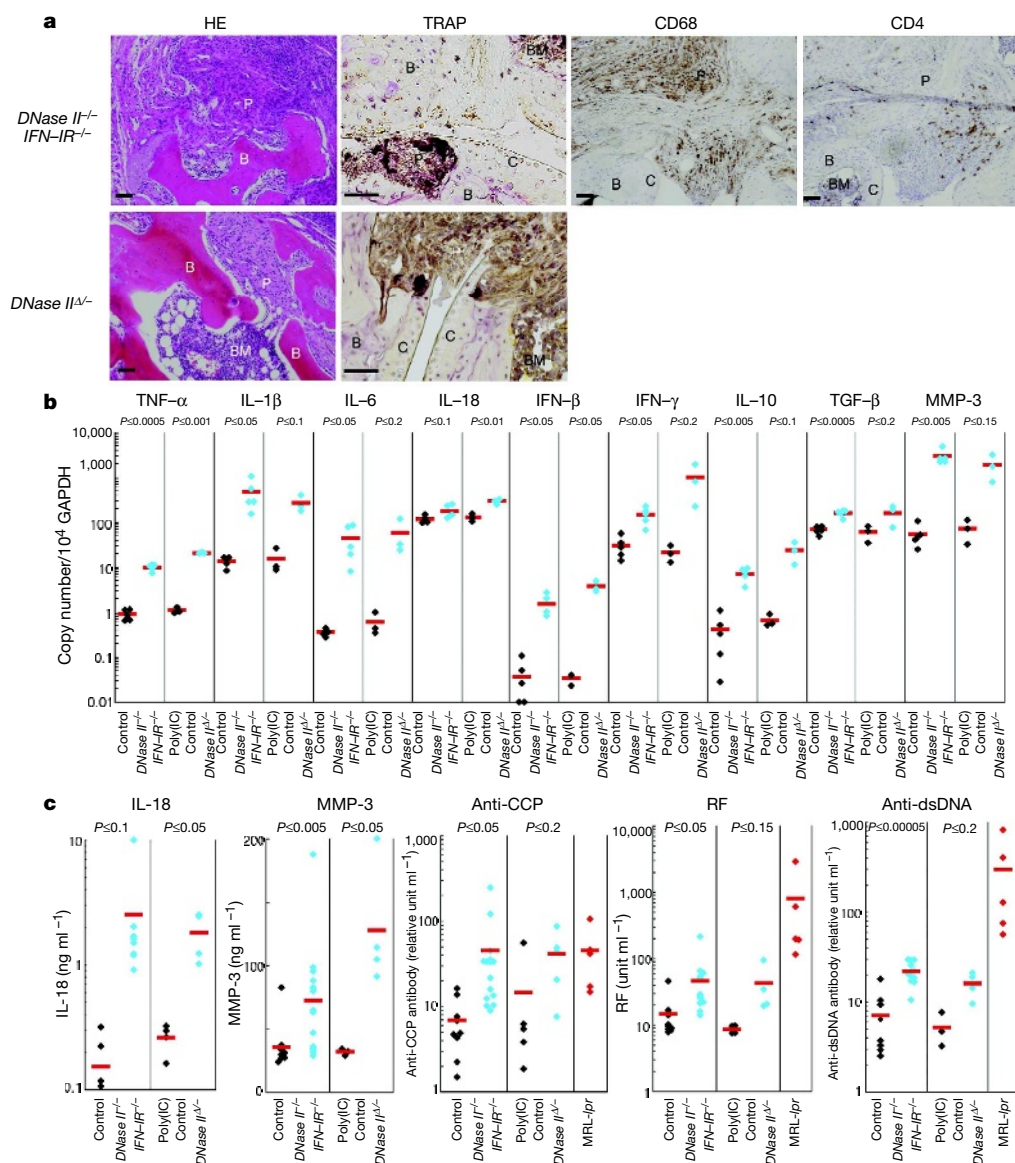


Figure 2 | Inflammatory arthritis in *DNase II^{-/-}IFN-IR^{-/-}* mice.

a, Joint sections from 8–16.5-month-old *DNase II^{-/-}IFN-IR^{-/-}* or *DNase II^{flox/-}* mice 10 months after treatment with poly(I)•poly(C) were stained with haematoxylin/eosin (HE), TRAP, anti-CD68 and anti-CD4. P, pannus; B, bone; C, cartilage; BM, bone marrow. Scale bar, 50 μ m. **b**, RNAs were prepared from the joints of 7.5–13.5-month-old control or *DNase II^{-/-}IFN-IR^{-/-}*, and those of poly(I)•poly(C)-treated *DNase II^{flox/+}* (poly(IC) control) or *DNase II^{flox/-}* mice. The mRNA levels for the indicated cytokines quantified by real-time PCR are expressed relative to mRNA for glyceraldehyde-3-phosphate dehydrogenase (GAPDH). **c**, The concentrations of IL-18, MMP-3, anti-CCP, rheumatoid factor (RF) and anti-dsDNA in the sera from 7.5–16.5-month-old control or *DNase II^{-/-}IFN-IR^{-/-}*, and from poly(I)•poly(C)-treated *DNase II^{flox/+}* or *DNase II^{flox/-}* mice are shown. Some parameters were determined for 3–7.5 month-old MRL-lpr mice. In **b** and **c**, horizontal red lines indicate the average values. Welch's *t*-test was used to test for difference, and *P* values are shown.

which is specifically observed in human patients with rheumatoid arthritis¹¹, and rheumatoid factor were also detected at high levels in the serum of *DNase II*^{-/-} *IFN-IR*^{-/-} and *DNase II*^{Δ/-} mice (Fig. 2c). The level of anti-double-stranded DNA (dsDNA) antibody in *DNase II*^{-/-} *IFN-IR*^{-/-} mice was about threefold that in wild-type mice, but this value was less than one-tenth of that observed in *lpr* mice, which develop an autoimmune disease similar to systemic lupus erythematosus¹².

DNase II^{-/-} *IFN-IR*^{-/-} and *DNase II*^{Δ/-} mice were slightly anaemic (hematocrit 40–45) and had clear growth retardation (data not shown). They developed splenomegaly from the age of 1 month, which became prominent as the mice aged; the weight of the spleen of 7.5–16.5-month-old *DNase II*^{-/-} *IFN-IR*^{-/-} mice was about fivefold that of control mice (Fig. 3a). Splenomegaly was due to enlargement of the red pulp. We and others have reported previously that fetal liver and thymus of *DNase II*^{-/-} mice contain many abnormal macrophages carrying undigested DNA from erythroid precursor cells or apoptotic cells^{3,4,13}. Similarly, numerous abnormal macrophages carrying DNA in lysosomes were found in bone marrow, spleen and other tissues of the adult *DNase II*^{-/-} *IFN-IR*^{-/-} and

DNase II^{Δ/-} mice (Fig. 3b). In addition, the serum of 4–6-week-old *DNase II*^{-/-} *IFN-IR*^{-/-} mice contained about 10 μg ml⁻¹ DNA, and this level decreased as the mice aged (Fig. 3c). *DNase II*^{Δ/-} mice also carried a low but significant level of DNA in their serum. It is likely that the undigested DNA leaked from the macrophages into the bloodstream.

Cytokines can affect the pathogenesis of human rheumatoid arthritis⁷; we therefore examined the involvement of cytokines in our model. In comparison with that of control mice (less than 20 pg ml⁻¹), the serum concentration of TNF-α in *DNase II*^{-/-} *IFN-IR*^{-/-} mice was higher (about 100 pg ml⁻¹) at 4–6 weeks of age (Fig. 3d), before the joints showed any apparent abnormality. A similar high concentration of TNF-α was observed in *DNase II*^{Δ/-} mice. In contrast, other cytokines, such as IL-6, IL-1β, IFN-γ, granulocyte/macrophage colony-stimulating factor and IL-4 were not undetectable in the serum of *DNase II* mutant mice. We previously reported that the fetal liver of *DNase II*^{-/-} embryos constitutively expresses various cytokines¹⁴. Real-time PCR indicated that the level of mRNA for TNF-α in the bone marrow of 1-month-old *DNase II*^{-/-} *IFN-IR*^{-/-} mice was twice that the level in age-matched control mice (Fig. 3e).

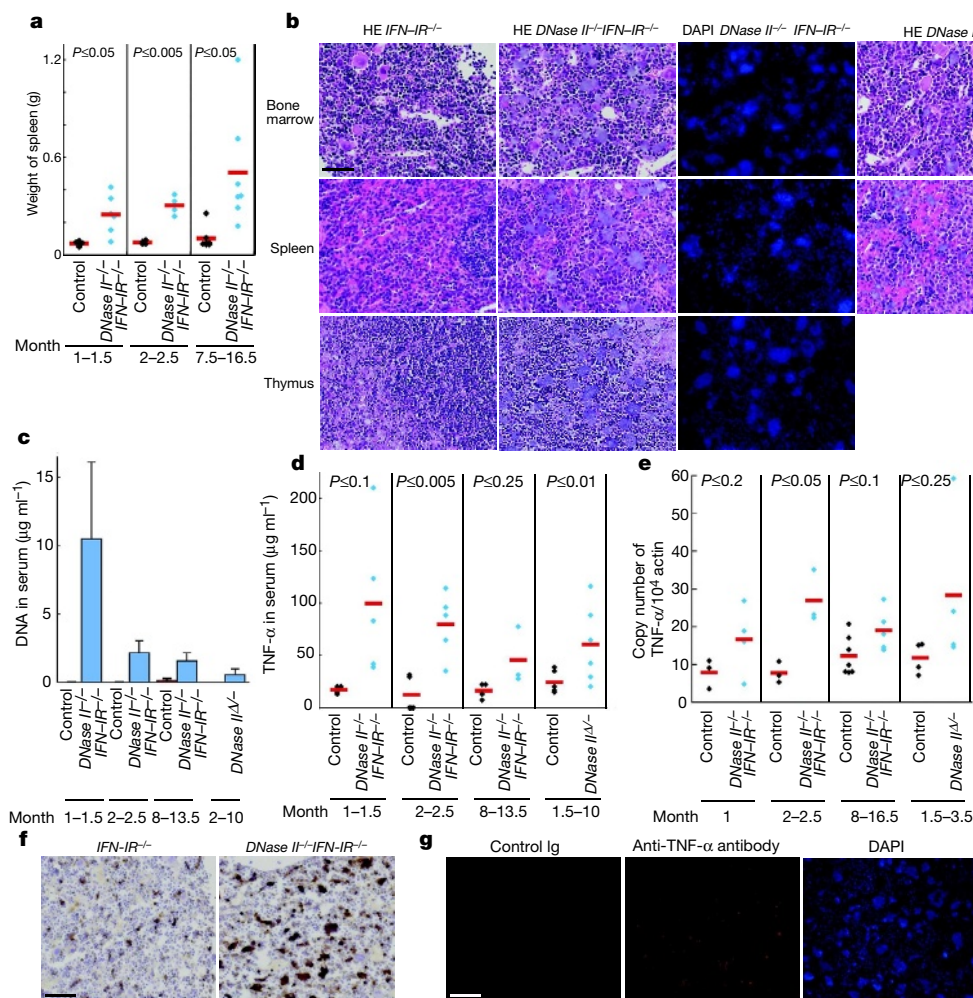


Figure 3 | Abnormal macrophages. **a**, The weight of spleens from control and *DNase II*^{-/-} *IFN-IR*^{-/-} (four to seven mice) at the indicated ages. **b**, Bone marrow, spleen and thymus of 4–8-week-old *IFN-IR*^{-/-} and *DNase II*^{-/-} *IFN-IR*^{-/-} and *DNase II*^{Δ/-} mice 10 months after treatment with poly(I)•poly(C) were stained with haematoxylin/eosin stain (HE) or 4',6-diamidino-2-phenylindole (DAPI). Scale bar, 50 μm. **c**, DNA concentration in the serum of control, *DNase II*^{-/-} *IFN-IR*^{-/-} and *DNase II*^{Δ/-} mice (two to six animals for each) at the indicated ages. Results are plotted as means and s.d. **d**, Serum TNF-α concentration of control and *DNase II*^{-/-} *IFN-IR*^{-/-}, and *DNase II*^{Δ/-} and *DNase II*^{fllox/+} littermates after

treatment with poly(I)•poly(C) (three to six mice for each). **e**, mRNA for TNF-α in the bone marrow of control and *DNase II*^{-/-} *IFN-IR*^{-/-}, and that of poly(I)•poly(C)-treated *DNase II*^{Δ/-} and *DNase II*^{fllox/+} littermates, quantified by real-time PCR, expressed relative to mRNA for β-actin. **f**, Bone marrow from 2-month-old *IFN-IR*^{-/-} and *DNase II*^{-/-} *IFN-IR*^{-/-} mice were stained with anti-CD68. Scale bar, 50 μm. **g**, Bone marrow from 1.5-month-old *DNase II*^{-/-} *IFN-IR*^{-/-} mice were stained with goat normal IgG or anti-mouse TNF-α. The DAPI staining profile is also shown. Scale bar, 50 μm. In **a**, **d** and **e**, horizontal red lines indicate the average values. Welch's *t*-test was used to test for difference, and *P* values are shown.

The level of mRNA for TNF- α had increased further in these mutant mice by the age of 2–2.5 months, when definitive erythropoiesis is still active. A similar high level of mRNA for TNF- α was found in the bone marrow of *DNase II*^{-/-} mice. Immunohistochemical analyses of the bone marrow (Fig. 3f) showed that the number of macrophages carrying the CD68 antigen in *DNase II*^{-/-}*IFN-IR*^{-/-} mice did not differ significantly from that in control mice. However, CD68 was more intensely labelled in *DNase II*^{-/-}*IFN-IR*^{-/-} mice, indicating that the macrophages carrying undigested DNA might have been activated¹⁵. Accordingly, TNF- α was detected in the cells carrying DNA (Fig. 3g).

To determine the effect of blocking TNF- α during the induction of arthritis, 4-week-old *DNase II*^{-/-}*IFN-IR*^{-/-} or *DNase II*^{+/-}*IFN-IR*^{-/-} littermates were given the neutralizing anti-TNF- α twice a week. *DNase II*^{-/-}*IFN-IR*^{-/-} mice that received phosphate-buffered saline (PBS) started to develop arthritis when they were 8–14 weeks old (Fig. 4a). Accordingly, the serum MMP-3 level increased significantly (Fig. 4b). In contrast, mice treated with anti-TNF- α did not develop arthritis at this age, and their serum MMP-3 level was comparable to or slightly higher than that of *DNase II*^{+/-}*IFN-IR*^{-/-} control mice. The expression of mRNA for TNF- α , IL-1 β , IL-6 and MMP-3 at the joints was also blocked by the anti-TNF- α (Fig. 4c).

We then examined the therapeutic effect of anti-TNF- α after *DNase II*^{-/-}*IFN-IR*^{-/-} mice had developed arthritis. Four *DNase II*^{-/-}*IFN-IR*^{-/-} mice that showed an arthritis clinical score of 2.5–6 were treated twice a week with anti-TNF- α , and the serum MMP-3 level, which decreases with anti-TNF- α therapy in humans¹⁶, was monitored. As shown in Fig. 4d, after 3 weeks of treatment the serum MMP-3 level had decreased to about one-third of its level before treatment. A decrease in the clinical score of the joints was

apparent in two mice. The foot joints were then removed from all the mice, and the cytokine mRNA levels were determined. The joints of *DNase II*^{-/-}*IFN-IR*^{-/-} mice receiving PBS carried 5–30-fold the levels of mRNA for TNF- α , IL-6, IL-1 β and MMP-3 than those of *DNase II*^{+/-}*IFN-IR*^{-/-} control mice, but the mRNA levels in the joints of *DNase II*^{-/-}*IFN-IR*^{-/-} mice treated with anti-TNF- α were significantly decreased (Fig. 4e).

Systemic expression of TNF- α in a transgenic mouse model causes the development of chronic polyarthritis, indicating that TNF- α can trigger the disease¹⁷. Moreover, human patients with rheumatoid arthritis respond well to anti-TNF- α therapy^{18,19}. However, the cause of the TNF- α gene expression in patients has not been established. We showed that macrophages carrying undigested DNA produce TNF- α , and that anti-TNF- α efficiently blocked the development of polyarthritis in *DNase II*-null mice. It is likely that TNF- α produced by *DNase II*^{-/-} macrophages triggered the pathogenesis by stimulating the growth of synovial cells and by activating cytokine genes in synovial tissues. More than 10¹⁰ red blood cells are generated in the human body every day, and the same number of nuclei are expelled from erythroid precursor cells and degraded by DNase II in bone marrow macrophages. The DNA of apoptotic cells, of which 10⁸–10⁹ are generated every day, is degraded by DNase II in macrophages. Our results indicate that a failure in this process can lead to the development of polyarthritis. Monocyte activation early in the onset of human rheumatoid arthritis has been noticed in some patients²⁰; this is consistent with our results and indicates that the activation of macrophages could be an initiator of the pathogenic cascade in human arthritis. In addition to anti-TNF- α therapy, human patients with rheumatoid arthritis respond to treatment

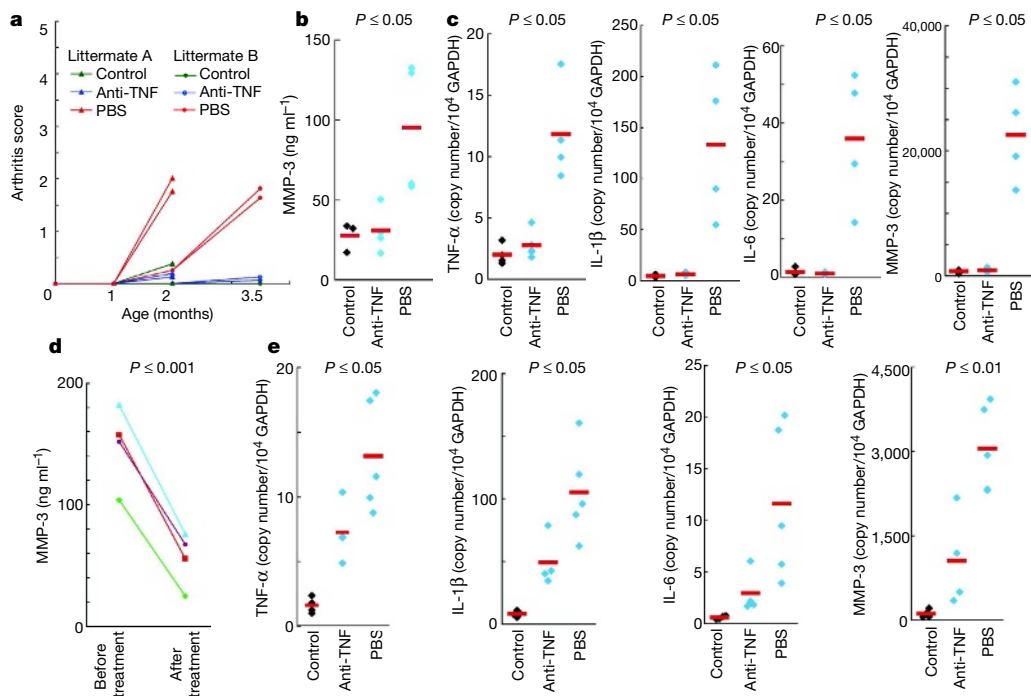


Figure 4 | Protective and therapeutic effect of anti-TNF- α .

a, *DNase II*^{-/-}*IFN-IR*^{-/-} (red and blue results) and *DNase II*^{+/-}*IFN-IR*^{-/-} (green results) mice were generated from two litters (shown as triangles and circles). At the age of 1 month, *DNase II*^{-/-}*IFN-IR*^{-/-} mice received anti-TNF- α (blue) or PBS (red) twice a week. The clinical score of polyarthritis was determined at the indicated ages. **b**, Serum MMP-3 level in 2-month-old *DNase II*^{+/-}*IFN-IR*^{-/-} (control), anti-TNF- α -treated or PBS-treated *DNase II*^{-/-}*IFN-IR*^{-/-} (four mice each). **c**, Levels of mRNA for TNF- α , IL-1 β , IL-6 and MMP-3 in the joints of the forefoot from 2–4-month-old *DNase II*^{+/-}*IFN-IR*^{-/-}, anti-TNF- α -treated or PBS-treated

DNase II^{-/-}*IFN-IR*^{-/-} mice expressed relative to mRNA for glyceraldehyde-3-phosphate dehydrogenase (GAPDH).

d, e, *DNase II*^{-/-}*IFN-IR*^{-/-} mice 5–13 months old showing a clinical score of 2.5–6.0 were treated twice a week with anti-TNF- α . **d**, The serum MMP-3 level was determined before and after the 3-week treatment. **e**, Levels of mRNA for TNF- α , IL-1 β , IL-6 and MMP-3 in the joints of *DNase II*^{+/-}*IFN-IR*^{-/-} (control) and anti-TNF- α -treated or untreated mice were determined after 6 weeks. In **b**, **c** and **e**, horizontal red lines indicate the average values. Welch's *t*-test was used to test for difference, and *P* values are shown.

targeting other pro-inflammatory cytokines, or lymphocytes^{21–23}. Whether these treatments are also effective for the polyarthritis in *DNase II*-null mice remains to be studied.

We found DNA in the serum of *DNase II*-deficient mice. Injection of bacterial or mitochondrial DNA into joints induces arthritis, which is mediated by TNF- α produced by macrophages in the joints^{24,25}. Immune complexes containing self-DNA activate rheumatoid-arthritis-specific B cells²⁶. In these cases, Toll-like receptor 9 (TLR9) is necessary for activating the TNF- α gene in macrophages and for producing rheumatoid factor in B cells. In contrast, deletion of the *TLR9* gene had no apparent effect on the development of arthritis in *DNase II*^{-/-} *IRF-IR*^{-/-} mice (K.K. and S.N., unpublished observations), indicating that DNA in the serum might not contribute to the pathogenesis directly. In contrast, we found that macrophages carrying undigested DNA were activated and produced TNF- α . Our results are consistent with the findings that the fetal liver macrophages in *DNase II*^{-/-} mice express IFN- β and TNF- α by a TLR-independent mechanism¹⁴ and that cytosolic DNA activates the innate immunity through a TLR-independent system^{27,28}. Further characterization of this TLR-independent pathway is needed to explain the molecular mechanism by which cells sense the DNA that has escaped from degradation and cause autoimmunity.

METHODS

Mice. Mice deficient in the *DNase II* and *IFN-IR* genes were described previously⁵. *E2a-Cre* (ref. 29) and *Mx1-Cre* transgenic mice⁶ were from the Jackson Laboratory. MRL-*lpr* mice were purchased from Japan SLC. The conditional DNase II targeting mice are described in Supplementary Methods. The anti-TNF- α monoclonal antibody (20 μ g per g body weight) was administered twice a week by intraperitoneal injection, and clinical parameters were assessed 3 days after the last injection.

Clinical assessment, and histology. Swelling of the forelimb and hindlimb joints was inspected manually and scored as follows: 0, no swelling; 1, mild swelling; 2, severe swelling or deformation of the limb or finger. The scores were summed, and a total score (maximum 8) was given to each mouse. The detailed procedures of MRI, radiography and histological analysis are described in Supplementary Methods.

Enzyme-linked immunosorbent assay (ELISA), real-time PCR and serum DNA. The serum levels of TNF- α , IL-18, rheumatoid factor and MMP-3 were determined with ELISA kits from Beckton Dickinson, MBL, Shibayagi and R&D Systems, respectively. Anti-dsDNA and anti-CCP were measured with a MESACUP DNA-II test kit (MBL) and a DIASTAT anti-CCP ELISA kit (Axis-Shield), except that peroxidase-conjugated goat antibody against mouse immunoglobulin (Cappel) was used as the detection antibody. The procedure for real-time PCR, the primers and the assay method for serum DNA are described in Supplementary Methods.

Received 27 July; accepted 14 September 2006.

- Nagata, S. DNA degradation in development and programmed cell death. *Annu. Rev. Immunol.* **23**, 853–875 (2005).
- Evans, C. J. & Aguilera, R. J. DNase II: genes, enzymes and function. *Gene* **322**, 1–15 (2003).
- Kawane, K. *et al.* Requirement of DNase II for definitive erythropoiesis in the mouse fetal liver. *Science* **292**, 1546–1549 (2001).
- Kawane, K. *et al.* Impaired thymic development in mouse embryos deficient in apoptotic DNA degradation. *Nature Immunol.* **4**, 138–144 (2003).
- Yoshida, H., Okabe, Y., Kawane, K., Fukuyama, H. & Nagata, S. Lethal anemia caused by interferon- β produced in mouse embryos carrying undigested DNA. *Nature Immunol.* **6**, 49–56 (2005).
- Kuhn, R., Schwenk, F., Aguet, M. & Rajewsky, K. Inducible gene targeting in mice. *Science* **269**, 1427–1429 (1995).
- Feldmann, M., Brennan, F. M. & Maini, R. N. Role of cytokines in rheumatoid arthritis. *Annu. Rev. Immunol.* **14**, 397–440 (1996).
- Firestein, G. S., Alvaro-Gracia, J. M. & Maki, R. Quantitative analysis of cytokine gene expression in rheumatoid arthritis. *J. Immunol.* **144**, 3347–3353 (1990).

- Saxne, T., Palladino, M. A. Jr, Heinegard, D., Talal, N. & Wollheim, F. A. Detection of tumor necrosis factor alpha but not tumor necrosis factor beta in rheumatoid arthritis synovial fluid and serum. *Arthritis Rheum.* **31**, 1041–1045 (1988).
- Manicourt, D. H., Fujimoto, N., Obata, K. & Thonar, E. J. Levels of circulating collagenase, stromelysin-1, and tissue inhibitor of matrix metalloproteinases 1 in patients with rheumatoid arthritis. Relationship to serum levels of antigenic keratan sulfate and systemic parameters of inflammation. *Arthritis Rheum.* **38**, 1031–1039 (1995).
- Schellekens, G. A., de Jong, B. A., van den Hoogen, F. H., van de Putte, L. B. & van Venrooij, W. J. Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies. *J. Clin. Invest.* **101**, 273–281 (1998).
- Cohen, P. L. & Eisenberg, R. A. *Lpr* and *gld*: single gene models of systemic autoimmunity and lymphoproliferative disease. *Annu. Rev. Immunol.* **9**, 243–269 (1991).
- Krieser, R. J. *et al.* Deoxyribonuclease IIa is required during the phagocytic phase of apoptosis and its loss causes lethality. *Cell Death Differ.* **9**, 956–962 (2002).
- Okabe, Y., Kawane, K., Akira, S., Taniguchi, T. & Nagata, S. Toll-like receptor-independent gene induction program activated by mammalian DNA escaped from apoptotic DNA degradation. *J. Exp. Med.* **202**, 1333–1339 (2005).
- Ramprasad, M. P., Terpstra, V., Kondratenko, N., Quehenberger, O. & Steinberg, D. Cell surface expression of mouse macrophage mannose receptor and human CD68 and their role as macrophage receptors for oxidized low density lipoprotein. *Proc. Natl Acad. Sci. USA* **93**, 14833–14838 (1996).
- Brennan, F. M. *et al.* Reduction of serum matrix metalloproteinase 1 and matrix metalloproteinase 3 in rheumatoid arthritis patients following anti-tumour necrosis factor-alpha (cA2) therapy. *Br. J. Rheumatol.* **36**, 643–650 (1997).
- Keffer, J. *et al.* Transgenic mice expressing human tumour necrosis factor: a predictive genetic model of arthritis. *EMBO J.* **10**, 4025–4031 (1991).
- Feldmann, M. & Maini, R. N. Anti-TNF alpha therapy of rheumatoid arthritis: what have we learned? *Annu. Rev. Immunol.* **19**, 163–196 (2001).
- Feldmann, M. Development of anti-TNF therapy for rheumatoid arthritis. *Nature Rev. Immunol.* **2**, 364–371 (2002).
- Fujii, I., Shingu, M. & Nobunaga, M. Monocyte activation in early onset rheumatoid arthritis. *Ann. Rheum. Dis.* **49**, 497–503 (1990).
- Edwards, J. C. & Cambridge, G. B-cell targeting in rheumatoid arthritis and other autoimmune diseases. *Nature Rev. Immunol.* **6**, 394–403 (2006).
- Bluestone, J. A., St Clair, E. W. & Turka, L. A. CTLA4lg: bridging the basic immunology with clinical application. *Immunity* **24**, 233–238 (2006).
- Yokota, S. *et al.* Therapeutic efficacy of humanized recombinant anti-interleukin-6 receptor antibody in children with systemic-onset juvenile idiopathic arthritis. *Arthritis Rheum.* **52**, 818–825 (2005).
- Deng, G. M., Nilsson, I. M., Verdrengh, M., Collins, L. V. & Tarkowski, A. Intracellularly localized bacterial DNA containing CpG motifs induces arthritis. *Nature Med.* **5**, 702–705 (1999).
- Collins, L. V., Hajizadeh, S., Holme, E., Jonsson, I. M. & Tarkowski, A. Endogenously oxidized mitochondrial DNA induces *in vivo* and *in vitro* inflammatory responses. *J. Leukoc. Biol.* **75**, 995–1000 (2004).
- Leadbetter, E. A. *et al.* Chromatin-IgG complexes activate B cells by dual engagement of IgM and Toll-like receptors. *Nature* **416**, 603–607 (2002).
- Stetson, D. B. & Medzhitov, R. Recognition of cytosolic DNA activates an IRF3-dependent innate immune response. *Immunity* **24**, 93–103 (2006).
- Ishii, K. J. *et al.* A Toll-like receptor-independent antiviral response induced by double-stranded B-form DNA. *Nature Immunol.* **7**, 40–48 (2006).
- Williams-Simons, L. & Westphal, H. EllaCre—utility of a general deleter strain. *Transgenic Res.* **8**, 53–54 (1999).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank K. Aozasa for pathological analysis of the mice, P. Quartier for critical reading of our manuscript, M. Nishikawa, H. Matsuda, T. Matsuki, A. Seiyama and T. Yanagida for advice and discussion, H. Fukuyama for help at the initial stage of this work, and M. Fujii and M. Harayama for secretarial assistance. This work was supported in part by Grants-in-Aid from the Ministry of Education, Science, Sports, and Culture in Japan.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to S.N. (nagata@genetic.med.osaka-u.ac.jp).

Amplification of histone genes by circular chromosome formation in *Saccharomyces cerevisiae*

Diana E. Libuda¹ & Fred Winston¹

Proper histone levels are critical for transcription, chromosome segregation, and other chromatin-mediated processes^{1–7}. In *Saccharomyces cerevisiae*, the histones H2A and H2B are encoded by two gene pairs, named *HTA1-HTB1* and *HTA2-HTB2* (ref. 8). Previous studies have demonstrated that when *HTA2-HTB2* is deleted, *HTA1-HTB1* dosage compensates at the transcriptional level^{4,9}. Here we show that a different mechanism of dosage compensation, at the level of gene copy number, can occur when *HTA1-HTB1* is deleted. In this case, *HTA2-HTB2* amplifies via creation of a new, small, circular chromosome. This duplication, which contains 39 kb of chromosome II, includes *HTA2-HTB2*, the histone H3-H4 locus *HHT1-HHF1*, a centromere and origins of replication. Formation of the new chromosome occurs by recombination between two Ty1 retrotransposon elements that flank this region. Following meiosis, recombination between these two particular Ty1 elements occurs at a greatly elevated level in *hta1-htb1Δ* mutants, suggesting that a decreased level of histones H2A and H2B specifically stimulates this amplification of histone genes. Our results demonstrate another mechanism by which histone gene dosage is controlled to maintain genomic integrity.

Although early genetic studies of *S. cerevisiae* genes encoding histones H2A and H2B clearly demonstrated that they are essential for growth^{10,11}, deletion of either of the *H2A-H2B* loci seemed to allow viability^{4,6}. In those studies, *hta2-htb2Δ* mutants were shown to grow normally, presumably owing to increased transcription of *HTA1-HTB1*^{4,9}. In contrast, *hta1-htb1Δ* mutants were shown to have several mutant phenotypes, including defects in transcription and chromatin structure^{3,4,6,12}. Notably, several unresolved mysteries have surrounded *hta1-htb1Δ* mutants. First, it was reported¹³ that some *hta1-htb1Δ* strains contain two copies of *HTA2-HTB2*, although this possible duplication was not characterized. Second, the *S. cerevisiae* deletion project reported that, whereas an *hta1Δ* mutant is viable, an *htb1Δ* mutant is inviable¹⁴. Finally, although *hta1-htb1Δ* mutants have been constructed in an S288C background, such mutants are inviable in W303 background (P. Kaufman and M. A. Osley, personal communication).

As an initial step in further characterizing *hta1-htb1Δ* mutants, we analysed strains previously suspected of containing an *HTA2-HTB2* duplication¹³. In this analysis we included the *hta1Δ* mutant from the *S. cerevisiae* deletion project¹⁴. To test for possible genomic changes, chromosomes were separated using contour-clamped homogeneous electric field (CHEF) gels and then probed by Southern analysis using an *HTA2-HTB2* probe (Fig. 1a). Our results show that for all *hta1-htb1Δ* strains and the *hta1Δ* strain, the *HTA2-HTB2* probe hybridized both to chromosome II, the normal location of *HTA2-HTB2*, and to a second prominent band (Fig. 1b). Thus, in these strains, *HTA2-HTB2* exists as a second copy that is not part of chromosome II.

To determine the size of the *HTA2-HTB2* amplification in *hta1-htb1Δ* mutants, we tested for amplification of flanking sequences. *HTA2-HTB2* is in a 33-kb region, flanked by Ty1 elements, that contains the centromere and 15 genes, including the histone *H3-H4* locus *HHT1-HHF1* (Fig. 1a; Supplementary Fig. 1). Our results show that sequences between the Ty1 elements display the same hybridization pattern as *HTA2-HTB2* in the *hta1-htb1Δ* strains (Fig. 1b, c). In contrast, sequences just outside the Ty1 elements were present in single copy in all strains tested (Fig. 1c). These results establish that the amplified region is delimited by the flanking Ty1 elements.

To determine whether all *hta1-htb1Δ* strains would contain the same amplification, additional *hta1-htb1Δ* strains were constructed by two methods. First, we deleted *HTA1-HTB1* in a diploid, sporulated the *HTA1-HTB1/hta1-htb1Δ* heterozygote, and dissected tetrads. Among 78 tetrads analysed, 149/156 *HTA1-HTB1* spores germinated and grew normally. In contrast, most *hta1-htb1Δ* spores failed to germinate, suggesting that this deletion causes inviability. Eventually, 20/156 *hta1-htb1Δ* spores formed colonies, appearing two days later than the wild-type colonies (Fig. 2a). (In a separate experiment, 70 *hta1-htb1Δ* spores that failed to form colonies were examined by microscopy and all failed to proceed past a single cell division (Supplementary Table 1; Supplementary Discussion).) When the viable *hta1-htb1Δ* strains were restreaked, they had a growth rate similar to that of the wild type (Supplementary Fig. 2a). To test the new viable *hta1-htb1Δ* strains for the amplification, 15 were examined by CHEF gel Southern analyses (Fig. 2b and data not shown) and all were shown to contain the amplification event observed in Fig. 1. This suggests that the amplification is required for the viability of *hta1-htb1Δ* strains and that it occurs at a high frequency in *hta1-htb1Δ* spores (see Supplementary Discussion). Additional experiments show that the *HTA2-HTB2* genes within the amplification are sufficient for viability in *hta1-htb1Δ* strains (Supplementary Fig. 3). Southern hybridizations and real-time polymerase chain reaction (PCR) determined that the amplification copy number is approximately one per cell (data not shown).

In a second approach, we constructed an *hta1-htb1Δ* haploid strain containing an *HTA1-HTB1 URA3* plasmid. Three independent cultures were plated on 5-fluoroorotic acid (5-FOA) medium to select for viable derivatives without the plasmid, thus generating new *hta1-htb1Δ* strains. Of the resistant (5-FOA^R) colonies generated, 6/32 examined contained the amplification (Supplementary Fig. 2b). The frequency of the amplification was determined to be 3×10^{-5} amplification events per cell (Supplementary Table 2; Supplementary Discussion). Of the *hta1-htb1Δ* strains without the amplification, ten were tested and shown to be chromosome II disomes; all ten had a severe growth defect (Supplementary Fig. 2). Altogether, these results show that *hta1-htb1Δ* strains require two copies of *HTA2-HTB2* for

¹Department of Genetics, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, Massachusetts 02115, USA.

viability. The remainder of our studies focused on the amplification event.

Several results suggested that the amplified version of *HTA2-HTB2* is circular and formed by homologous recombination between the Ty1 elements (Supplementary Discussion). To test for a circle, three experiments were performed. First, PCR was performed with primers designed to amplify a product only if the Ty1-Ty1 recombination on chromosome II occurred. Our results show that, in contrast to wild-type strains, all *hta1-htb1Δ* strains tested produced the PCR product expected from the Ty1-Ty1 recombinant (Fig. 3a). Second, Southern analyses of restriction-enzyme-digested genomic DNA were probed for fragments that would exist in only the recombinants. Our results demonstrate a novel restriction fragment with a size consistent with a 39 kb circular chromosome formed by recombination between the Ty1 elements (Supplementary Fig. 4). Third, homothallic switching (HO) endonuclease was used to cut the amplified DNA *in vivo*. If the amplified DNA is circular, then linearization by HO endonuclease digestion would cause a large shift in its migration in CHEF gels. Therefore, *hta1-htb1Δ* strains were constructed with a HO endonuclease site integrated near the *HTA2-HTB2* locus and

with a plasmid containing the *HO* gene under the control of the *GAL1* promoter. Our results show that upon *HO* induction, the migration of the amplified DNA shifts markedly, consistent with the linearization of a 39 kb circular chromosome (Fig. 3b; compare lanes 15 and 16). Collectively, these results provide strong evidence that the amplified *HTA2-HTB2* locus is a novel circular chromosome.

Two experiments were performed to examine the role of Ty1 elements in the amplification. First, to test whether both Ty1 elements on chromosome II are required for the amplification, *HTA1-HTB1/hta1-htb1Δ* diploids that were homozygous for deletion of one of the two Ty1 elements were constructed. For the deletion of *YBLWTy1-1*, 37 tetrads were analysed and only 4/74 *hta1-htb1Δ* spores formed colonies. For the deletion of *YBRWTy1-2*, 46 tetrads were analysed and only 1/92 *hta1-htb1Δ* spores grew. All five of these strains grew extremely slowly, and none of them contained the amplification (Fig. 4a). Genetic analysis suggested that these strains are disomic for chromosome II. The conclusion that the amplification requires both Ty1 elements is also supported by experiments with *S. cerevisiae* W303. Unlike the S288C background, deletion of *HTA1-HTB1*

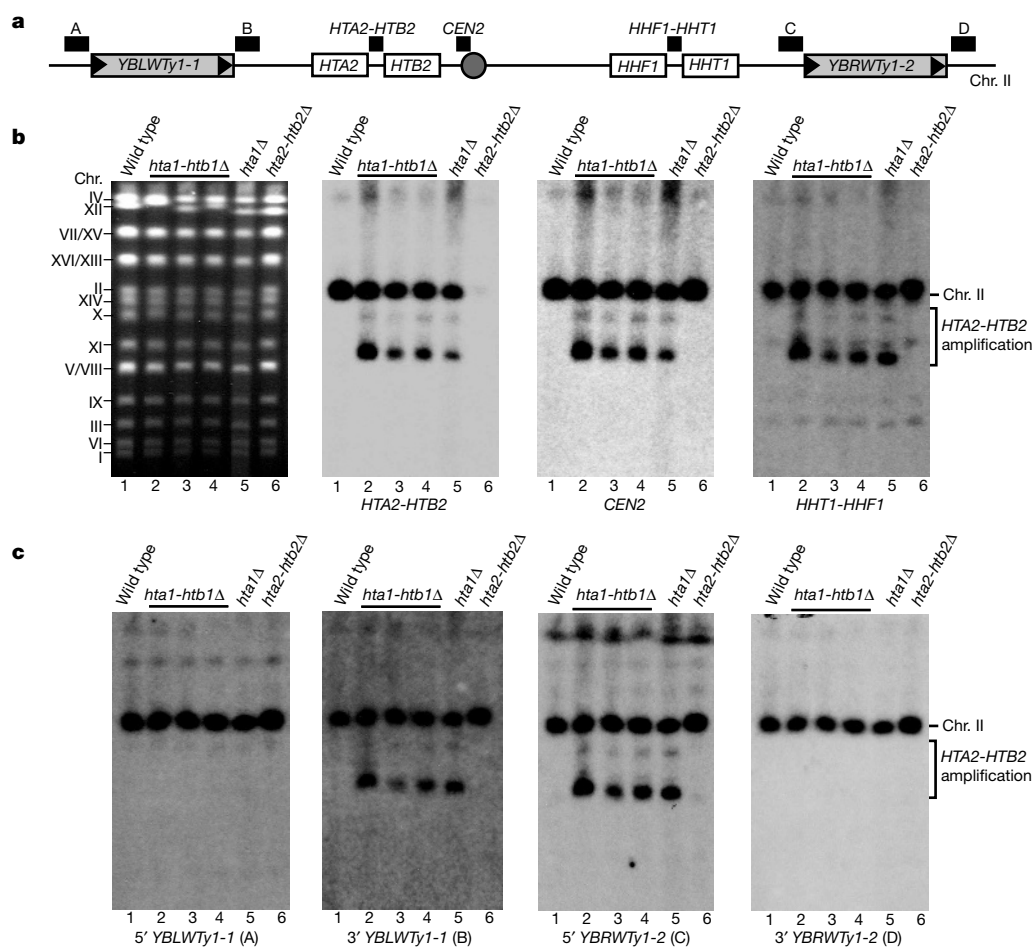


Figure 1 | Characterization of *HTA2-HTB2* amplification in *hta1-htb1Δ*. **a**, A simplified genome map of the *HTA2-HTB2* region on chromosome II. The white boxes represent the *HTA2-HTB2* and *HHT1-HHF1* histone loci. The grey boxes represent the Ty1 elements, *YBLWTy1-1* and *YBLWTy1-2*, with the long terminal repeats (known as LTRs or δ elements) shown as black triangles. The grey circle represents the centromere. The probes used in **b** and **c** are represented as black bars above the map in **a**; the probe used for each Southern analysis is indicated below the blots in **b** and **c**. A complete map of the region is shown in Supplementary Fig. 1. **b**, Separated chromosomes from *S. cerevisiae* strains were analysed on a CHEF gel. Each panel shows a wild-type strain (lane 1), three *hta1-htb1Δ* strains (lanes 2–4), the *hta1Δ* strain from the deletion project (lane 5), and an *hta2-htb2Δ* strain (lane 6).

The left panel is an ethidium-bromide-stained gel showing the positions of the chromosomes. The altered migration of chromosome XII is a consequence of the rDNA repeats, which are known to vary in number²⁶. The other three panels show analysis of the same gel by Southern hybridization analysis. **c**, Chromosomes from the same strains were examined as in **b** using probes just outside (probes A and D) and within (probes B and C) the region flanked by *YBLWTy1-1* and *YBLWTy1-2*. The *hta1Δ* strain from the deletion set has the identical pattern to our *hta1-htb1Δ* strains. We assume that during construction of the deletion set, the *hta1Δ* strain acquired the amplification and the *htb1Δ* strain did not. Chr., chromosome.

causes inviability in all *hta1-htb1Δ* spores in the W303 background (P. Kaufman and M. A. Osley, personal communication). Our analysis of W303 has shown that only a solo δ element exists in place of *YBRWty1-2* (Supplementary Fig. 5). Taken together, our results strongly suggest that each of the Ty1 elements on chromosome II is required for the amplification event.

Second, to test whether the amplification can occur if the Ty1 elements are replaced by a different sequence of similar length, we replaced each Ty1 with linearized plasmid Ylp5 (ref. 15). Ylp5 is 5.5 kb long (compared with 5.9 kb for Ty1) and contains no homology to Ty1 elements. To measure the recombination frequency, *HTA1-HTB1/(hta1-htb1Δ)* diploids homozygous for the Ty1 replacements with Ylp5 were sporulated. From 115 tetrads, 44/230 *hta1-htb1Δ* spores (19.1%) contained the amplification, a frequency comparable to that observed for amplification with the Ty1 elements. Both CHEF gels (Fig. 4b) and PCR (data not shown) confirmed that the amplification was an extrachromosomal circle. Thus, Ty1 elements are not required for the amplification, suggesting that any homologous sequence of adequate length is sufficient.

The amplification event that we have described occurs at a greatly elevated frequency in *hta1-htb1Δ* mutants following meiosis, compared with the frequencies of previously studied Ty-Ty recombination events^{16–18}. This high frequency might reflect a genome-wide elevation of recombination in *hta1-htb1Δ* mutants due to reduced levels of histones H2A and H2B. Alternatively, the high frequency of recombination may be specific to this pair of Ty1 elements. To distinguish between these possibilities, we measured the recombination frequency for two additional pairs of adjacent Ty1 elements under the same conditions in which the histone gene amplification occurs. As a

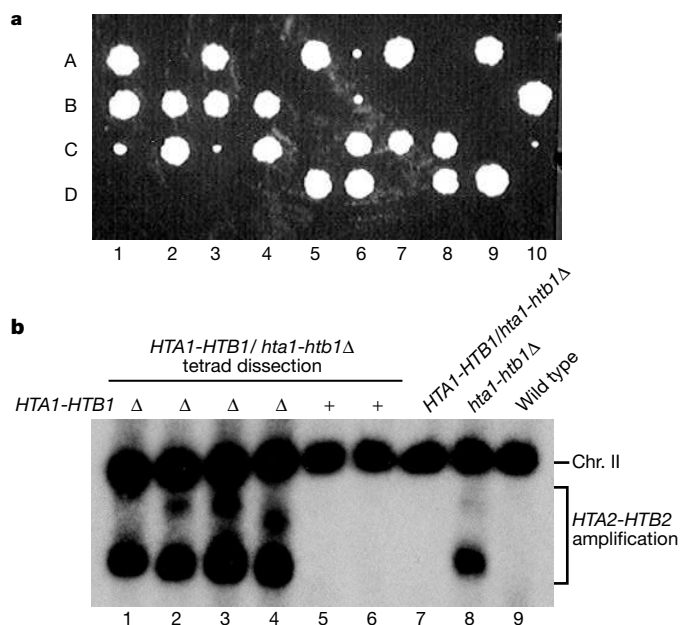


Figure 2 | Analysis of newly constructed *hta1-htb1Δ* strains. **a**, Strains from *HTA1-HTB1/hta1-htb1Δ* tetrad dissection. Shown is a representative tetrad dissection plate. Each column shows growth of the four progeny from a single tetrad. The large colonies are wild type and the small colonies are *hta1-htb1Δ*. The apparent slow growth of the *hta1-htb1Δ* colonies is caused by slow germination (see Supplementary Discussion). Most spores that failed to germinate, for example, tetrad 1 spore D, were also *hta1-htb1Δ*. **b**, Four *hta1-htb1Δ* progeny (lanes 1–4) and two wild-type progeny (lanes 5 and 6) generated from the *HTA1-HTB1/(hta1-htb1Δ)* tetrad dissection were analysed by a CHEF gel Southern analysis probed with *HTA2-HTB2*. Controls include the *HTA1-HTB1/hta1-htb1Δ* parent diploid (lane 7), an *hta1-htb1Δ* haploid (lane 8) and a wild-type strain (lane 9). All strains were tested with the same probes as Fig. 1 and shown to contain the same amplified region (data not shown).

control, we first showed that each of these pairs of Ty1 elements has normal levels of mitotic recombination (Supplementary Table 3). Our results (Fig. 4c) show that the level of Ty1-Ty1 recombination that gives rise to the amplification is significantly higher than for either of the other Ty1-Ty1 recombination events following meiosis, demonstrating that there is not a genome-wide elevation of recombination. Rather, the recombination event that forms the amplification appears to be enhanced specifically at this locus in *hta1-htb1Δ* spores.

How can a high level of recombination between two specific Ty elements be explained? Indeed, previous studies have shown that Ty elements are generally cold for meiotic recombination and, in fact, suppress the activity of nearby recombination hotspots¹⁹. Several factors have been suggested to control meiotic recombination levels, including chromatin structure and histone modifications²⁰. In

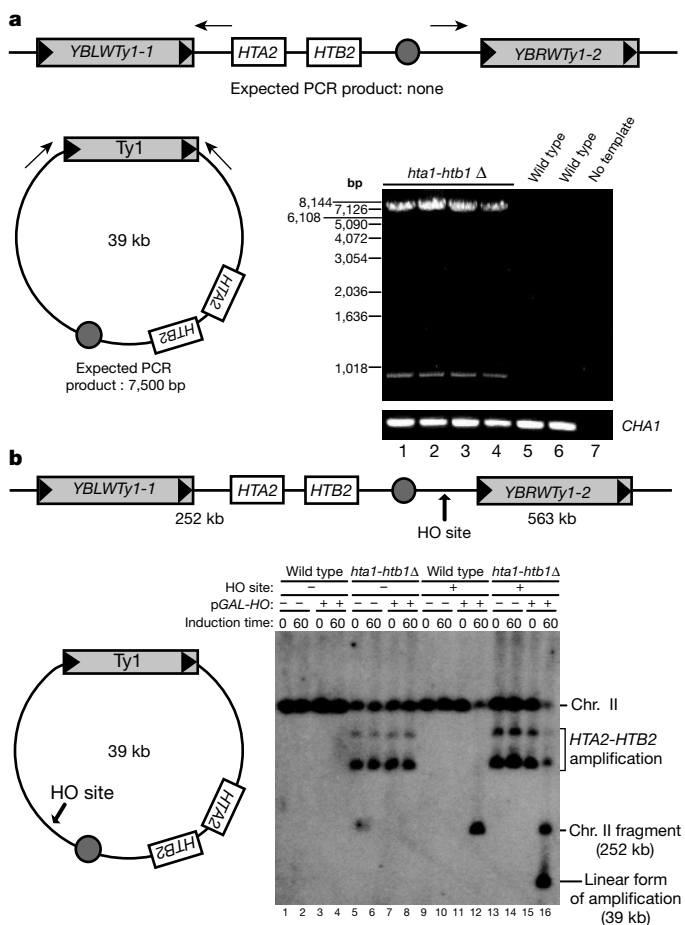


Figure 3 | Tests for the presence of a 39 kb circular chromosome. **a**, PCR test for recombination between *YBLWty1-1* and *YBLWty1-2*. The left side shows an abbreviated view of the *HTA2-HTB2* region of chromosome II, before and after recombination between the two Ty1 elements. The arrows indicate the PCR primers designed to amplify a product only if the recombination event occurred. The right side shows an ethidium-bromide-stained gel of the PCR products. Lanes 1–4 are four independently derived *hta1-htb1Δ* strains containing the amplification. The presence of the 7,500 bp PCR product is consistent with the predicted recombination event. The less intense 930 bp band is probably due to crossover PCR between partly synthesized products. *CHA1* is a positive control for the PCR. **b**, Cleavage of the amplification by HO endonuclease. The left diagram indicates the location of the HO endonuclease site on chromosome II and the amplification. The right side shows an *HTA2-HTB2*-probed CHEF gel Southern analysis of wild-type and *hta1-htb1Δ* strains, with and without the HO site, before and after HO induction. The 252 kb band in lanes 12 and 16 is a fragment of chromosome II produced by HO cleavage. The 39 kb band in lane 16 indicates a linearized form of the amplification.

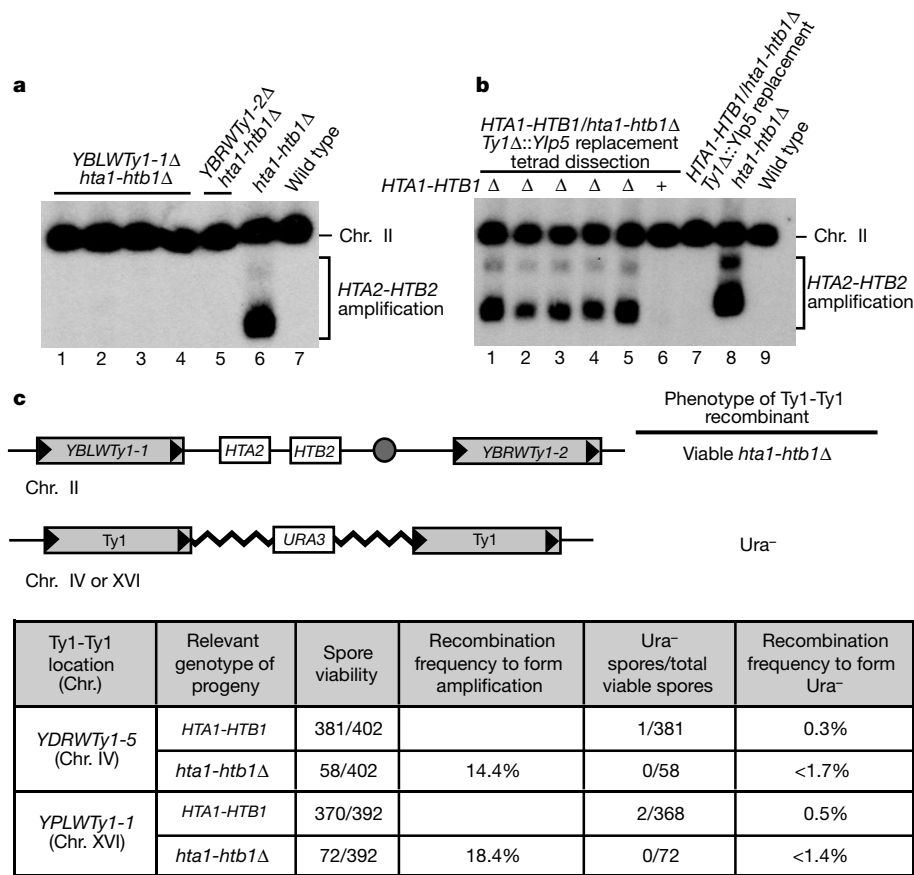


Figure 4 | Characterization of the recombination event to amplify histone genes. **a**, Analysis of Ty1 deletion mutants. *HTA1-HTB1/hta1-htb1Δ* diploids that are homozygous for deletion of either *YBLWTy1-1* or *YBLWTy1-2* were sporulated and *hta1-htb1Δ* survivors were analysed for the amplification. Shown is a CHEF gel Southern analysis probed with *HTA2-HTB2*. *YBLWTy1-1Δ hta1-htb1Δ* strains (lanes 1–4) and the *YBLWTy1-2Δ hta1-htb1Δ* strain (lane 5) do not contain the *HTA2-HTB2* amplification. Controls are an *hta1-htb1Δ* mutant with normal Ty1 elements (lane 6) and a wild-type strain (lane 7). **b**, Replacement of the Ty1 elements. *HTA1-HTB1/hta1-htb1Δ* diploids in which the Ty1 elements *YBLWTy1-1* and *YBRWTy1-2* were each replaced with Ylp5, were sporulated and dissected. Surviving *hta1-htb1Δ* strains (lanes 1–5) and wild-type progeny (lane 6) were analysed as described for **a**. Controls include the parent diploid (lane 7), an *hta1-htb1Δ* haploid (lane 8), and a wild-type strain (lane 9). By PCR, the *HTA2-HTB2* amplification in the Ty1 replacement *hta1-htb1Δ* strains was confirmed to be a circular chromosome (data not shown). Note that owing to its slightly smaller size, the band for the Ylp5-containing amplification migrates faster than the Ty1-containing amplification (compare lanes 1–5 with lane 8). **c**, Frequency of Ty1-Ty1

recombination in wild-type and *hta1-htb1Δ* strains. The diagram depicts the Ty1 flanked region of chromosome II with the *HTA2-HTB2* locus and the Ty1-*URA3*-Ty1 configuration on chromosome IV or XVI. In the Ty1-*URA3*-Ty1 constructs, the Ty1 elements are 17 kb apart, compared with 33 kb for the Ty1 elements on chromosome II. The expected phenotypes of recombinants at both classes of loci are indicated. *HTA1-HTB1/hta1-htb1Δ* diploid strains were constructed that are homozygous for either Ty1-*URA3*-Ty1 configuration. The diploids were sporulated and analysed for the phenotypes of recombinants at either locus. The table presents the results of the Ty1-Ty1 recombination assay. The amplification frequency was found to be 14–18% in *hta1-htb1Δ*. In contrast, none of the viable *hta1-htb1Δ* progeny lost the *URA3* marker, demonstrating that the *HTA2-HTB2* amplification event occurs at a significantly higher frequency than the Ty1-*URA3*-Ty1 recombination event. All three pairs of Ty1 elements are capable of normal levels of recombination during mitotic growth (Supplementary Table 3). In a separate experiment, the *HTA2-HTB2* amplification frequency in wild-type strains following meiosis was found to be close to 0% (see Supplementary Methods).

addition, earlier analysis of *hta1-htb1Δ* mutants demonstrated that they have localized, rather than general, effects on chromatin structure³. Thus, when H2A and H2B levels are reduced, as in *hta1-htb1Δ* spores, there may be a specific local alteration of chromatin structure to create a hotspot for recombination, resulting in the observed amplification frequency. Conceivably, the recombination frequency could also be affected by transcription levels, as has been suggested for the *S. cerevisiae* recombination enhancer that controls the directionality of mating-type switching^{21,22}. However, our results suggest that Ty1 transcription is not a critical factor (see Supplementary Discussion).

Ty elements have been suggested to mediate genomic rearrangements under selective pressure²³. Our work has identified a previously unknown mechanism by which Ty elements enable histone genes to dosage-compensate in response to reduced histone levels. To our knowledge, this is the first example of a natural role for Ty

elements that is dependent upon their specific genomic position. This amplification mechanism would also allow the transient modulation of histone levels in wild-type cells in response to the need for altered histone levels; the amplification could occur when more histones are required, and the circular chromosome could be easily lost when this requirement ends. The amplification described here suggests that similar transposon-related mechanisms may serve in adaptive gene amplification in other organisms, including humans, where 45% of the genome consists of transposons^{24,25}.

METHODS

Details of strains, plasmids, media, CHEF gel analysis, Southern blot hybridization analysis, PCR across Ty elements, and HO endonuclease experiments can be found in Supplementary Information.

Ty1-Ty1 recombination assay. Ty1-Ty1 recombination markers were constructed to determine the level of Ty1-Ty1 recombination in wild-type and

hta1-htb1Δ strains. To do this, plasmid B155, which contains a complete Ty1 and *URA3*, was used to transform FY26. Integrants resulting from recombination between the Ty1 element on the plasmid and a Ty1 element in the genome result in two Ty1 elements flanking 17 kb of plasmid sequences including *URA3*. The integration location was mapped using CHEF gels, Southern blots and PCR analysis. Two different integrants were chosen for further analysis: in FY2526, B155 integrated at *YPLWTy1-1* on chromosome XVI, and in FY2523, B155 integrated at *YDRWTy1-5* on chromosome IV. By a series of crosses (FY2526 × FY2523 and FY2523 × FY2524) and selection on 5-fluoroanthranilic acid (5-FAA) media to lose the pDL2 plasmid, *HTA1-HTB1/hta1-htb1Δ* diploid strains, homozygous for either of the Ty1-Ty1 recombination markers, were constructed. For Fig. 4c, the resultant diploids were then sporulated and tetrads dissected to assess the stability of the Ty1-Ty1 recombination marker in both wild-type and *hta1-htb1Δ* spores. If the frequency of Ty1-Ty1 recombination throughout the genome is elevated to the same frequency of the Ty1-Ty1 recombination event that generates the amplification (14–18%), then 14–18% of *hta1-htb1Δ* strains with the amplification would be expected to be *Ura*⁺.

Received 17 July; accepted 6 September 2006.

- Meeks-Wagner, D. & Hartwell, L. H. Normal stoichiometry of histone dimer sets is necessary for high fidelity of mitotic chromosome transmission. *Cell* **44**, 43–52 (1986).
- Gunjan, A. & Verreault, A. A. Rad53 kinase-dependent surveillance mechanism that regulates histone protein levels in *S. cerevisiae*. *Cell* **115**, 537–549 (2003).
- Norris, D., Dunn, B. & Osley, M. A. The effect of histone gene deletions on chromatin structure in *Saccharomyces cerevisiae*. *Science* **242**, 759–761 (1988).
- Norris, D. & Osley, M. A. The two gene pairs encoding H2A and H2B play different roles in the *Saccharomyces cerevisiae* life cycle. *Mol. Cell. Biol.* **7**, 3473–3481 (1987).
- Han, M., Chang, M., Kim, U. J. & Grunstein, M. Histone H2B repression causes cell-cycle-specific arrest in yeast: effects on chromosomal segregation, replication, and transcription. *Cell* **48**, 589–597 (1987).
- Clark-Adams, C. D., Norris, D., Osley, M. A., Fassler, J. S. & Winston, F. Changes in histone gene dosage alter transcription in yeast. *Genes Dev.* **2**, 150–159 (1988).
- Han, M. & Grunstein, M. Nucleosome loss activates yeast downstream promoters *in vivo*. *Cell* **55**, 1137–1145 (1988).
- Hereford, L., Fahrner, K., Woolford, J. Jr, Rosbash, M. & Kaback, D. B. Isolation of yeast histone genes H2A and H2B. *Cell* **18**, 1261–1271 (1979).
- Moran, L., Norris, D. & Osley, M. A. A yeast H2A–H2B promoter can be regulated by changes in histone gene copy number. *Genes Dev.* **4**, 752–763 (1990).
- Rykowski, M. C., Wallis, J. W., Choe, J. & Grunstein, M. Histone H2B subtypes are dispensable during the yeast cell cycle. *Cell* **25**, 477–487 (1981).
- Kolodrubetz, D., Rykowski, M. C. & Grunstein, M. Histone H2A subtypes associate interchangeably *in vivo* with histone H2B subtypes. *Proc. Natl Acad. Sci. USA* **79**, 7814–7818 (1982).
- Hirschhorn, J. N., Brown, S. A., Clark, C. D. & Winston, F. Evidence that SNF2/SWI2 and SNF5 activate transcription in yeast by altering chromatin structure. *Genes Dev.* **6**, 2288–2298 (1992).
- Hirschhorn, J. N., Bortvin, A. L., Ricupero-Hovasse, S. L. & Winston, F. A new class of histone H2A mutations in *Saccharomyces cerevisiae* causes specific transcriptional defects *in vivo*. *Mol. Cell. Biol.* **15**, 1999–2009 (1995).
- Giaever, G. *et al.* Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* **418**, 387–391 (2002).
- Struhl, K., Stinchcomb, D. T., Scherer, S. & Davis, R. W. High-frequency transformation of yeast: autonomous replication of hybrid DNA molecules. *Proc. Natl Acad. Sci. USA* **76**, 1035–1039 (1979).
- Kupiec, M. & Petes, T. D. Allelic and ectopic recombination between Ty elements in yeast. *Genetics* **119**, 549–559 (1988).
- Kupiec, M. & Petes, T. D. Meiotic recombination between repeated transposable elements in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **8**, 2942–2954 (1988).
- Roeder, G. S. Unequal crossing-over between yeast transposable elements. *Mol. Gen. Genet.* **190**, 117–121 (1983).
- Ben-Aroya, S., Mieczkowski, P. A., Petes, T. D. & Kupiec, M. The compact chromatin structure of a Ty repeated sequence suppresses recombination hotspot activity in *Saccharomyces cerevisiae*. *Mol. Cell* **15**, 221–231 (2004).
- Petes, T. D. Meiotic recombination hot spots and cold spots. *Nature Rev. Genet.* **2**, 360–369 (2001).
- Wu, X. & Haber, J. E. A 700 bp cis-acting region controls mating-type dependent recombination along the entire left arm of yeast chromosome III. *Cell* **87**, 277–285 (1996).
- Ercan, S., Reese, J. C., Workman, J. L. & Simpson, R. T. Yeast recombination enhancer is stimulated by transcription activation. *Mol. Cell. Biol.* **25**, 7976–7987 (2005).
- Zeyl, C. Capturing the adaptive mutation in yeast. *Res. Microbiol.* **155**, 217–223 (2004).
- Venter, J. C. *et al.* The sequence of the human genome. *Science* **291**, 1304–1351 (2001).
- Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
- Schwartz, D. C. & Cantor, C. R. Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell* **37**, 67–75 (1984).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank A. Dudley and D. Helmlinger for helpful comments on the manuscript. We also thank J. Haber and J.-A. Kim for the suggestion of and advice on the HO experiment, and V. Dror for instruction about CHEF gels. We are grateful to A. Gabriel, P. Kaufman, M. A. Osley and T. Petes for sharing unpublished results and for discussions. We thank J. Hirschhorn, whose observations led to this project. This work was supported by a grant from the National Institutes of Health to F.W. and by a National Science Foundation Graduate Fellowship to D.E.L.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to F.W. (winston@genetics.med.harvard.edu).

LETTERS

Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation

Yijun Qi^{1†}, Xingyue He^{1,2}, Xiu-Jie Wang³, Oleksiy Kohany⁴, Jerzy Jurka⁴ & Gregory J. Hannon¹

DNA methylation has important functions in stable, transcriptional gene silencing, immobilization of transposable elements and genome organization¹. In *Arabidopsis*, DNA methylation can be induced by double-stranded RNA through the RNA interference (RNAi) pathway, a response known as RNA-directed DNA methylation². This requires a specialized set of RNAi components, including ARGONAUTE4 (AGO4)^{3–6}. Here we show that AGO4 binds to small RNAs including small interfering RNAs (siRNAs) originating from transposable and repetitive elements, and cleaves target RNA transcripts. Single mutations in the Asp-Asp-His catalytic motif of AGO4 do not affect siRNA-binding activity but abolish its catalytic potential. siRNA accumulation and non-CpG DNA methylation at some loci require the catalytic activity of AGO4, whereas others are less dependent on this activity. Our results are consistent with a model in which AGO4 can function at target loci through two distinct and separable mechanisms. First, AGO4 can recruit components that signal DNA methylation in a manner independent of its catalytic activity. Second, AGO4 catalytic activity can be crucial for the generation of secondary siRNAs that reinforce its repressive effects.

RNA-directed DNA methylation (RdDM) involves a class of siRNAs about 24 nucleotides (nt) in length, which are believed to confer sequence specificity on the process. *Arabidopsis* has evolved a set of RNAi components that are specialized for RdDM, including Dicer-like 3 (DCL3), RNA-dependent RNA polymerase 2 (RDR2), RNA polymerase IV and ARGONAUTE4 (AGO4)^{3,4,7,8}. Mutations in these proteins can lead to decreased accumulation of siRNAs, decreased AGO4 stability⁹, and decreased DNA methylation at many endogenous loci including transposons and repetitive elements^{3–8,10}. It is highly probable that RNA polymerase IV, RDR2 and DCL3 are components of the siRNA biogenesis machinery. AGO4 is the prime candidate for the component of the effector complex that directs DNA methylation as guided by siRNAs.

We therefore tested whether AGO4 exists in a complex with siRNAs *in vivo*. We generated an *Arabidopsis Landsberg erecta* (Lar) transgenic line expressing a tandem affinity purification (TAP)-tagged AGO4 protein. This protein was recovered from whole plant extracts¹¹ and its associated RNAs were examined by SYBR-gold staining. TAP-AGO4 was associated predominantly with small RNAs about 24 nt in length (Fig. 1a). Parallel examination of AGO1 complexes showed prominent small RNAs about 21 nt in length^{12,13} (Fig. 1a). Northern blotting revealed the binding of AGO4 to siRNAs originating from known transposons, specifically *AtSN1* (ref. 14), *AtMu1* (ref. 15) and a repeated sequence, *MEA-ISR* (ref. 16) (Fig. 1b), whose methylation is known to be controlled by RNAi⁴.

To obtain a more complete catalogue of the small RNAs associated with AGO4, we cloned and subjected them to sequencing as described¹⁷. For comparison, the total small RNA population, ranging from 18 to 28 nt, and small RNAs associated with AGO1 were also sequenced. In all, 74,390 sequences were obtained for the whole population (total), and 55,497 and 193,167 sequences were obtained from AGO4 and AGO1 complexes, respectively. As a quality control, we mapped all candidate small RNA sequences to the *Arabidopsis* genome and found that 51,294 (total), 21,198 (AGO4) and 152,088 (AGO1) sequences perfectly matched at least one location (Supplementary Table S1). Only small RNAs passing this quality assessment were used in further analyses.

Small RNA sequences from the total population showed two discernible peaks at 21 and 24 nt. Most AGO4-associated RNAs were 23–24 nt in length, whereas AGO1-associated small RNAs were almost exclusively 21 nt (Fig. 1c); 10,058 (47%) of the AGO4-associated RNAs matched repetitive sequences in the genome, whereas only 15% of total and 3% of the AGO1-associated RNAs were repeat-derived (Supplementary Table S2). Such repeats comprise 17 of the 18 different types documented in Repbase¹⁸ (Supplementary Table S3), with the top 30 families accounting for more than 50% of all matches (Supplementary Table S4). Genomic matches to AGO4-associated small RNAs were particularly dense in pericentromeric regions, reflecting their high concentration of repetitive sequences (Fig. 1d, and Supplementary Figs S1–S4). Although we cannot unambiguously determine whether a given small RNA was derived from a particular repeat copy, previous studies indicate that AGO4 complexes can act *in trans* to direct RdDM at matching loci^{5,6}. One must therefore consider these plots as reflecting sites of possible action rather than sites of possible origin.

In all, 91% of AGO1-associated small RNAs matched known microRNAs (miRNAs) (Supplementary Table S5), a result consistent with the demonstrated role of AGO1 in miRNA-mediated control of plant development^{12,13,19}. Only a small fraction of AGO4-associated RNAs (3%) matched microRNAs. However, a subset of the microRNAs found in AGO4 complexes showed preferential association with this protein (Supplementary Fig. S5). This implies selectivity in how individual microRNAs are processed and passed to specific AGO-containing RNA-induced silencing complexes (RISCs). Moreover, it implies that some microRNAs might need to act in concert with AGO4, perhaps in the nucleus, rather than with AGO1 in the cytoplasm. Many AGO4-associated small RNAs were also derived from the sense or antisense strand of genes, pseudogenes and intergenic regions (Supplementary Table S5), indicating that AGO4 might also have a previously unrecognized general role in regulating gene expression.

¹Cold Spring Harbor Laboratory, Watson School of Biological Sciences and Howard Hughes Medical Institute, 1 Bungtown Road, Cold Spring Harbor, New York 11724, USA. ²Program in Genetics, Stony Brook University, Stony Brook, New York 11794, USA. ³State Key Laboratory of Plant Genomics, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Datun Road, 100101 Beijing, China. ⁴Genetic Information Research Institute, 1925 Landings Drive, Mountain View, California 94043, USA. [†]Present address: National Institute of Biological Sciences, No. 7 Science Park Road, Zhongguanchun Life Science Park, 102206 Beijing, China.

ATREP2 (a *Helitron*-like DNA transposon with dispersed repeats in the genome) and *SIMPLEHAT2* (a DNA transposon) both matched abundant small RNAs that are associated with AGO4 (Supplementary Fig. S6a). Both the levels of these small RNAs and non-CpG methylation at these loci were reduced by genetic lesions in the AGO4 pathway (in *RDR2*, *DCL3* and *AGO4*; Supplementary Fig. S6b, c). Taken together, our data indicate that siRNAs associated with AGO4 direct it to target loci, where it can promote non-CpG (CpNpG and CpHpH) methylation.

It is not known how siRNAs act at target loci to direct RdDM. Opposing classes of models involve either pairing between an siRNA and a target DNA (RNA–DNA recognition) or pairing between an siRNA and a nascent RNA transcript (RNA–RNA recognition)²⁰. However, methylation of *PHABULOSA* can be directed by a miR165/166 target site that crosses an exon–exon junction, strongly supporting an RNA–RNA recognition model²¹. A crucial question is whether this siRNA–RNA pairing leads to transcript cleavage and whether such cleavage functions in silencing²².

Alignment of *Arabidopsis* AGO4 with known Slicers, human Ago2 (ref. 23) and *Arabidopsis* AGO1 (refs 12, 13), revealed the presence of the catalytic Asp–Asp–His (DDH) triad in all three proteins²⁴ (Fig. 2a). To test its catalytic potential directly, we incubated TAP-purified AGO4 with an RNA transcript containing a sequence complementary to identified *ATREP2* siRNAs. The target RNA was cleaved by wild-type AGO4 protein (Fig. 2b) but not by mutants containing changes in essential catalytic residues (D660A, D742A and H874A substitutions, respectively; referred to hereafter as DDH mutants), despite similar expression levels and siRNA-binding capacity (Fig. 2c,

and Supplementary Fig. S7). We could similarly detect target cleavage on incubation of AGO4 complexes with targets for two AGO4-interacting miRNAs, miR172 and miR390 (Supplementary Fig. S8).

To determine whether catalysis was important for RdDM, we turned to a system of epialleles that could be tracked through an obvious visual phenotype. *SUPERMAN* is required for proper floral development; when its activity is decreased, plants show an increased number of stamens (an average of ten in comparison with the normal six) and incompletely fused carpels (*SUP* phenotype). In addition to genetic mutants, there are also *SUP* epialleles (Clark Kent or *clk*)²⁵. When the *clk-3* epiallele is placed in an AGO4-null background (*clk-3/ago4-1*, obtained from the *Arabidopsis* Biological Resource Center), most plants are wild-type (Fig. 3a), although some (20–30%) do retain the *SUP* phenotype. This indicates that there might be complex regulation of the locus and that *SUP* silencing in a subset of plants might be maintained in the absence of AGO4. However, when the *clk-3/ago4-1* plants were viewed as a population, CpNpG and CpHpH methylation at *SUP* were decreased substantially (Fig. 3b, and Supplementary Table S7) in comparison with *clk-st* (a stabilized Clark Kent allele in the presence of wild-type AGO4)⁴.

AGO4 or each DDH mutant was expressed under the control of the AGO4 promoter in *clk-3/ago4-1* plants. Pooled samples from about 30 primary transformants (*T*₁ generation) showed that all proteins were expressed at similar levels (Supplementary Fig. S7). Essentially all *T*₁ plants transformed with wild-type AGO4 complemented *ago4-1* and displayed the *SUP* floral phenotype, whereas those with the empty vector did not (Fig. 3a). Intriguingly, all *T*₁ plants transformed with AGO4 DDH mutants also recovered the *SUP* phenotype

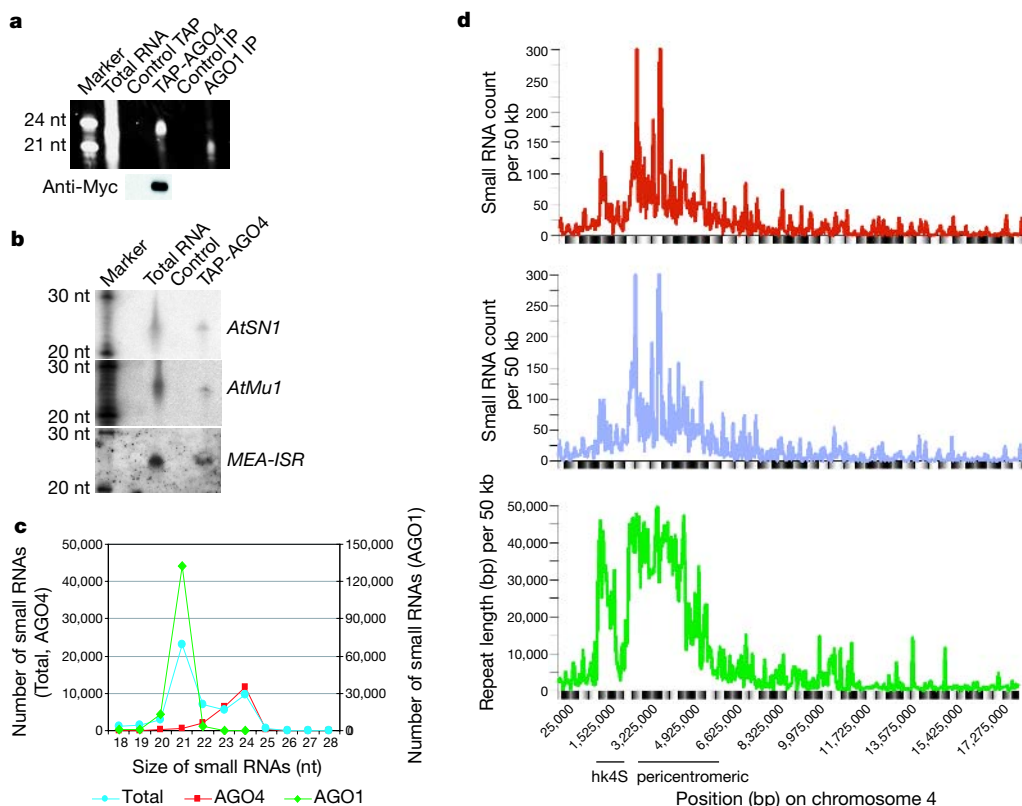


Figure 1 | A catalogue of AGO4-associated small RNAs. **a**, SYBR-gold was used to reveal small RNAs in total *Arabidopsis* RNA, TAP-AGO4 complexes, AGO1 complexes and control purifications (upper panel, as indicated). Western blotting with anti-Myc antibody detected AGO4 in the TAP-AGO4 purification but not in the control purification (lower panel). IP, immunoprecipitation. **b**, Northern blotting was used to detect small RNAs in total RNA, TAP-AGO4 complex and control purifications with the indicated probes. Radioactive RNAs of known sizes were included as markers. **c**, Size distribution of total (cyan), AGO4-associated (red) and

AGO1-associated (green) small RNAs. The sets of redundant small RNAs were used to generate a histogram quantifying the number of sequences obtained for each size class. **d**, Chromosome-wide density analysis of the AGO4-associated small RNAs on chromosome 4. The density of small RNAs with perfect matches in the direct strand (upper panel) and the complementary strand (middle panel), and the density of repeats (presented as the total length of repeats (bp); lower panel) in a 50-kilobase sliding window, are plotted. The positions of the pericentromeric region and the heterochromatic knob *hk4S* are marked.

(Fig. 3a). Bisulphite sequencing of about 30 pooled seedlings of each genotype showed that non-CpG methylation was restored to approximately normal (*clk-st*) levels (Fig. 3b, and Supplementary Table S7). Although the regulation of SUP is likely to be complex, an intact AGO4 pathway reinforced silencing at the locus in a manner that did not depend on siRNA-directed RNA cleavage.

Examination of additional loci revealed a more complex picture (Fig. 4a, and Supplementary Table S7). At *AtMu1*, non-CpG methylation decreased in *ago4-1* plants⁴ and was fully rescued by AGO4 or its DDH mutants. However, at three other loci, namely *MEA-ISR*, *ATREP2* and *SIMPLEHAT2*, wild-type AGO4 restored non-CpG methylation to normal levels, but the DDH mutants showed greatly decreased potency. For example, at *MEA-ISR*, introduction of AGO4 into *ago4-1* mutants led to a 3.2-fold increase in CpNpG and a 20.9-fold increase in CpHpH methylation. The DDH mutants had from no effect to a roughly 2.1-fold increase in CpNpG and a 2.2-fold to 4.2-fold increase in CpHpH methylation. Thus, the requirement in RdDM for AGO4 catalytic potential varies with the locus.

We next probed the correlation between non-CpG methylation, siRNA production and the effect of inactivating the AGO4 catalytic site. At *AtMu1*, a locus where DDH mutants complemented

methylation defects efficiently, loss of AGO4 had no effect in itself on the abundance of *AtMu1* siRNAs, and these species were not increased on expression of any AGO4 variant (Fig. 4b). We were not able to examine the effect of DDH mutations on the accumulation of SUP siRNAs because these were below detection limits, as reported previously⁴. For the three repetitive elements for which the wild-type and DDH mutants showed differential effects, another pattern was observed. In all cases, siRNAs were substantially decreased in *ago4-1* mutant plants. Expression of wild-type AGO4 generally restored siRNA levels, but the DDH mutants had much less pronounced effects (Fig. 4b). For example, in the *MEA-ISR* locus, siRNAs in *ago4-1* mutants decreased to 18% of wild-type levels. Wild-type AGO4 restored this to 87% of normal, whereas the DDH mutants rescued siRNAs to only 27–38%. In comparison, siRNA02 was not decreased in *ago4-1* as described previously⁶, and the introduction of either catalytic or non-catalytic AGO4 had no effect on overall siRNA02 abundance.

Thus, the catalytic activity of AGO4 was important both for efficient siRNA production and for non-CpG methylation at some loci. At others, where AGO4 loss had little effect on overall siRNA levels, AGO4 loss could still affect non-CpG methylation. However, in these cases a lack of catalytic potential was of little importance to the ability of ectopically expressed AGO4 to restore non-CpG methylation. All of these conclusions were based on multiple independent T₁ transgenesis studies and bisulphite experiments (up to five each), and all results were confirmed with two individual T₂ transgenic lines for

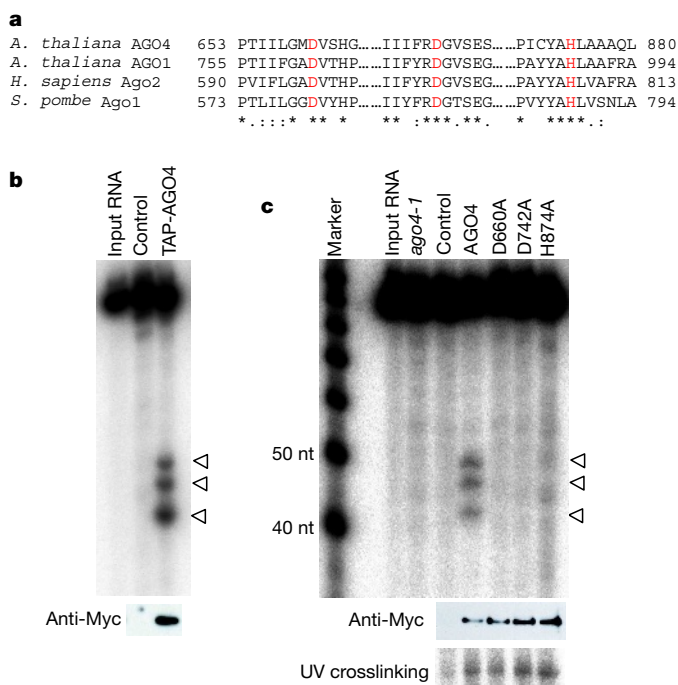


Figure 2 | AGO4 is a slicer. **a**, A partial alignment of the PIWI domains of *Arabidopsis* AGO4, AGO1, human Ago2 and *Schizosaccharomyces pombe* Ago1 is shown. The residues forming the catalytic DDH motif are shown in red. The degree of similarity is indicated under the alignment: asterisk indicates identity in all sequences; colon indicates conservative substitutions, and full point indicates semi-conservative substitutions. The starting and ending positions of the sequences are as labelled. **b**, A ³²P-labelled synthetic target RNA containing recognition sites for cloned ATREP2 siRNAs was incubated with TAP-purified AGO4 or a control purification. Positions of 5' products of cleavages guided by three endogenous ATREP2 siRNAs are indicated by the arrows (upper panel). Western blotting with anti-Myc antibody detected AGO4 in the TAP-AGO4 purification but not in the control purification (lower panel). **c**, The synthetic target RNA was incubated with immunopurified Myc-tagged AGO4 wild-type and DDH mutants (top panel). Decade RNA markers are shown for reference. Western blotting with an anti-Myc antibody detected AGO4 in the immunoprecipitates but not in the control purification (middle panel). AGO4 and control immunoprecipitates, as indicated, were incubated with single-stranded ³²P-labelled 24-nt siRNAs bearing photoreactive dT residues at the two 3' positions¹². Mixtures were irradiated with ultraviolet (UV) as described in Methods (bottom panel).

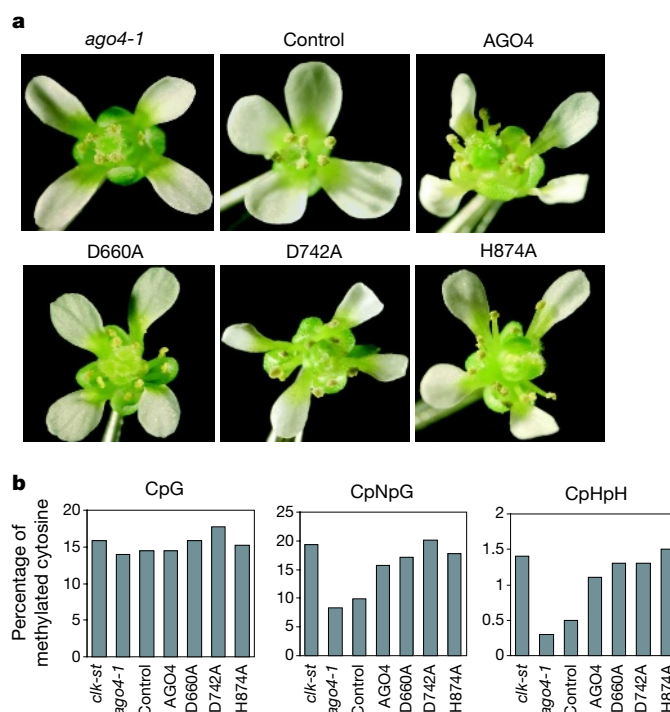


Figure 3 | Slicer activity is not required for non-CpG methylation and silencing at the SUP locus. **a**, Representative flowers from parental *clk-3/ago4-1* and plants of the same genotype transformed with vectors as indicated. About 70–80% of the *clk-3/ago4-1* plants and plants transformed with empty vector have wild-type flowers with six stamens and two fused carpels, with the remainder having the SUP phenotype (about ten stamens and three incompletely fused carpels; see the text). Essentially all flowers from plants transformed with AGO4 and DDH mutants display the SUP phenotype. **b**, CpG (left), CpNpG (centre) and CpHpH (right) methylation of the SUP gene was analysed by bisulphite sequencing of genomic DNA prepared from pooled T₁ seedlings. Data from two complete biological replicates were combined. The methylation level is shown by the percentage of methylated cytosine in all sequenced clones. The data in Supplementary Table S7 were used to generate the histograms.

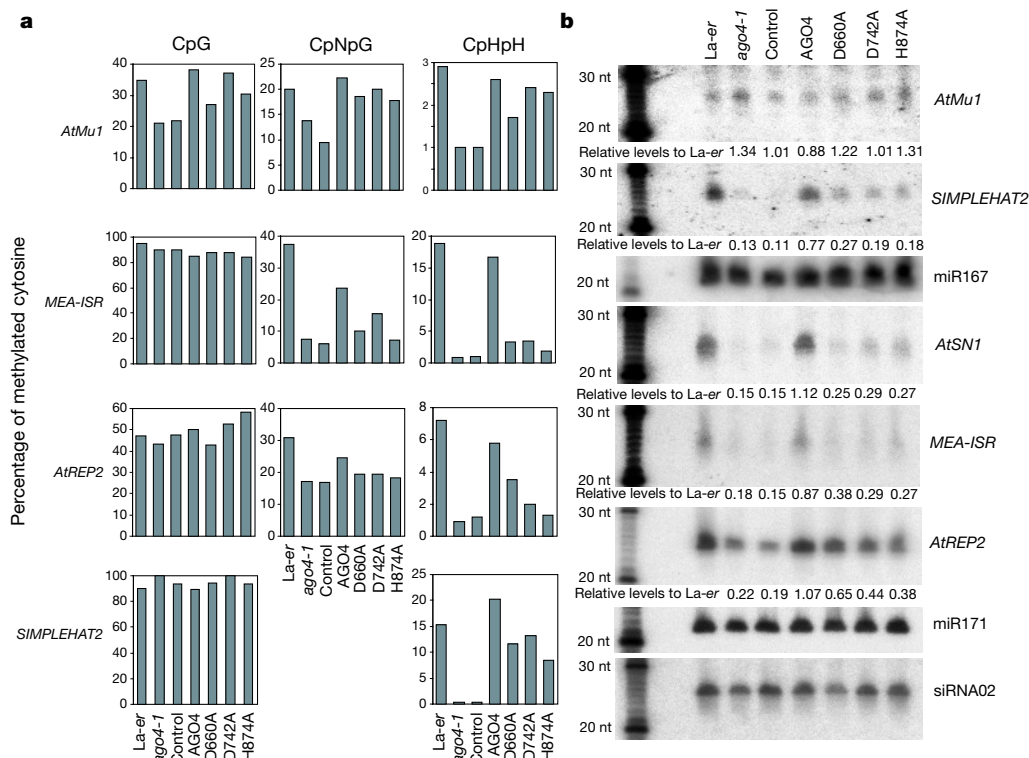


Figure 4 | Distinct effects of Slicer activity on DNA methylation and siRNA accumulation at endogenous repeats. **a**, CpG (left), CpNpG (middle) and CpHpH (right) methylation at *AtMu1*, *MEA-ISR*, *ATREP2* and *SIMPLEHAT2* loci were analysed by bisulphite sequencing. Data from two complete biological replicates were combined. Methylation levels are shown by the percentage of methylated cytosines in all sequenced clones. Data from Supplementary Table S7 were used to generate the histograms. **b**, Northern

blotting was used to analyse siRNAs derived from *AtMu1*, *SIMPLEHAT2*, *AtSN1*, *MEA-ISR*, *ATREP2* and *siRNA02* in RNA prepared from the indicated pooled T₁ plants. miR167 and miR171 were used as loading controls. The siRNA signals were normalized relative to miR167 (for *AtMu1* and *SIMPLEHAT2*) or miR171 (for other siRNAs), and the relative levels were calculated by comparison with those in *La-er* RNA (arbitrarily set to 1.00).

each construct (Supplementary Figs S9 and S10, and Supplementary Table S8).

Our data indicate that AGO4 can have two distinct functions in RdDM. First, AGO4 can direct chromatin remodelling factors to a target locus, probably through interactions between siRNAs and a nascent transcript. For this process, the catalytic activity of AGO4 is not required. Thus, given a source of siRNAs, the non-catalytic activity of AGO4 would be sufficient to sustain methylation and repression. Second, AGO4 has a function in which catalysis is required for efficient siRNA production. Cleavage may trigger the recruitment of RDR2-containing complexes to synthesize a double-stranded RNA using the cleaved transcript as template, with subsequent processing by DCL3 producing secondary siRNAs. This is reminiscent of the production of *Arabidopsis trans*-acting siRNAs, which is initiated by the cleavage of their precursor RNA by a miRNA-directed RISC²⁶. Other, currently mysterious, mechanisms must also promote siRNA production from heterochromatic loci, because in this model an existing siRNA or miRNA would be required to initiate the cycle. siRNA accumulation and non-CpG DNA methylation of *AtMu1*, and by inference at *SUP*, are much less dependent on the catalytic activity of AGO4. This could simply indicate that another AGO protein functions redundantly with AGO4 at these sites⁴. Indeed, it was shown that DNA methylation at *AtMu1* is also controlled by AGO1 (ref. 10). It also remains possible that *AtMu1* and *SUP* might represent a subset of AGO4-dependent loci in which the role of siRNAs is less important, particularly considering that *SUP* siRNAs have yet to be detected.

Our results reveal a potentially general property of Argonaute proteins. A single Argonaute may simultaneously serve as a catalytic engine of RNA cleavage and as a flexible platform for the assembly of multiprotein complexes that trigger cleavage-independent repression.

For AGO4, both of these functions act within a single silencing pathway to contribute to the management of repetitive sequences in the *Arabidopsis* genome.

METHODS

Complete details of methods used are given in Supplementary Information.

Construction of *Arabidopsis transgenic lines*. Transgenic plants were constructed with the use of a floral dip method²⁷ to transfer plasmids encoding a TAP-tagged or Myc-tagged wild-type or mutant AGO4 under the control of the AGO4 promoter.

Purification of AGO4 and AGO1 complexes. AGO1 and AGO4 complexes were purified from total extracts essentially as described previously^{11,12}.

Small RNA cloning. Small RNAs were gel isolated and used to create libraries for 454 Life Sciences sequencing as described previously¹⁷.

Slicer activity and siRNA-binding assays. TAP or immunopurified AGO4 complexes were incubated with target RNAs to examine the Slicer activity of AGO4 essentially as described¹². The siRNA-binding assay was performed as described¹².

Bisulphite sequencing. Genomic DNAs were isolated from 3–4-week-old seedlings with the DNeasy Plant Mini kit (Qiagen). The EZ DNA Methylation-Gold kit (ZYMO Research) was used for bisulphite treatment of genomic DNA in accordance with the manufacturer's instructions.

Received 22 June; accepted 4 September 2006.

Published online 24 September 2006.

- Chan, S. W., Henderson, I. R. & Jacobsen, S. E. Gardening the genome: DNA methylation in *Arabidopsis thaliana*. *Nature Rev. Genet.* **6**, 351–360 (2005).
- Matzke, M. A. & Birchler, J. A. RNAi-mediated pathways in the nucleus. *Nature Rev. Genet.* **6**, 24–35 (2005).
- Xie, Z. *et al.* Genetic and functional diversification of small RNA pathways in plants. *PLoS Biol.* **2**, e104 (2004).
- Zilberman, D., Cao, X. & Jacobsen, S. E. ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. *Science* **299**, 716–719 (2003).

5. Chan, S. W. *et al.* RNA silencing genes control *de novo* DNA methylation. *Science* **303**, 1336 (2004).
6. Zilberman, D. *et al.* Role of *Arabidopsis* ARGONAUTE4 in RNA-directed DNA methylation triggered by inverted repeats. *Curr. Biol.* **14**, 1214–1220 (2004).
7. Herr, A. J., Jensen, M. B., Dalmay, T. & Baulcombe, D. C. RNA polymerase IV directs silencing of endogenous DNA. *Science* **308**, 118–120 (2005).
8. Onodera, Y. *et al.* Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* **120**, 613–622 (2005).
9. Li, C. *et al.* An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in *Arabidopsis thaliana*. *Cell* **126**, 93–106 (2006).
10. Lippman, Z., May, B., Yordan, C., Singer, T. & Martienssen, R. Distinct mechanisms determine transposon inheritance and methylation via small interfering RNA and histone modification. *PLoS Biol.* **1**, e67 (2003).
11. Rohila, J. S., Chen, M., Cerny, R. & Fromm, M. E. Improved tandem affinity purification tag and methods for isolation of protein heterocomplexes from plants. *Plant J.* **38**, 172–181 (2004).
12. Qi, Y., Denli, A. M. & Hannon, G. J. Biochemical specialization within *Arabidopsis* RNA silencing pathways. *Mol. Cell* **19**, 421–428 (2005).
13. Baumberger, N. & Baulcombe, D. C. *Arabidopsis* ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. *Proc. Natl Acad. Sci. USA* **102**, 11928–11933 (2005).
14. Hamilton, A., Voinnet, O., Chappell, L. & Baulcombe, D. Two classes of short interfering RNA in RNA silencing. *EMBO J.* **21**, 4671–4679 (2002).
15. Singer, T., Yordan, C. & Martienssen, R. A. Robertson's Mutator transposons in *A. thaliana* are regulated by the chromatin-remodeling gene *Decrease in DNA Methylation (DDM1)*. *Genes Dev.* **15**, 591–602 (2001).
16. Cao, X. & Jacobsen, S. E. Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. *Proc. Natl Acad. Sci. USA* **99** (Suppl. 4), 16491–16498 (2002).
17. Girard, A., Sachidanandam, R., Hannon, G. & Carmell, M. A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature* **442**, 199–202 (2006).
18. Jurka, J. *et al.* Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet. Genome Res.* **110**, 462–467 (2005).
19. Vaucheret, H., Vazquez, F., Crete, P. & Bartel, D. P. The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes Dev.* **18**, 1187–1197 (2004).
20. Grewal, S. I. & Moazed, D. Heterochromatin and epigenetic control of gene expression. *Science* **301**, 798–802 (2003).
21. Bao, N., Lye, K. W. & Barton, M. K. MicroRNA binding sites in *Arabidopsis* class III HD-ZIP mRNAs are required for methylation of the template chromosome. *Dev. Cell* **7**, 653–662 (2004).
22. Qi, Y. & Hannon, G. J. Uncovering RNAi mechanisms in plants: biochemistry enters the fray. *FEBS Lett.* **579**, 5899–5903 (2005).
23. Liu, J. *et al.* Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305**, 1437–1441 (2004).
24. Rivas, F. V. *et al.* Purified Argonaute2 and an siRNA form recombinant human RISC. *Nature Struct. Mol. Biol.* **12**, 340–349 (2005).
25. Jacobsen, S. & Meyerowitz, E. Hypermethylated SUPERMAN epigenetic alleles in *Arabidopsis*. *Science* **277**, 1100–1103 (1997).
26. Allen, E., Xie, Z., Gustafson, A. M. & Carrington, J. C. MicroRNA-directed phasing during *trans*-acting siRNA biogenesis in plants. *Cell* **121**, 207–221 (2005).
27. Clough, S. J. & Bent, A. F. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743 (1998).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank S. Ossowski and D. Weigel for help with the analysis of AGO1-associated RNAs; R. Sachidanandam for bioinformatics assistance; T. Mulligan for assistance with *Arabidopsis* culture; and members of the Hannon laboratory for discussion. This work was supported in part by grants from the NSF and the NIH (G.J.H.). G.J.H. is an Investigator of the Howard Hughes Medical Institute.

Author Information Sequences referred to in this study have been deposited in the GenBank database under accession numbers DQ927324–DQ972825. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to G.J.H. (hannon@cshl.edu).

CORRIGENDUM

doi:10.1038/nature05152

The DNA sequence and biological annotation of human chromosome 1

S. G. Gregory, K. F. Barlow, K. E. McLay, R. Kaul, D. Swarbreck, A. Dunham, C. E. Scott, K. L. Howe, K. Woodfine, C. C. A. Spencer, M. C. Jones, C. Gillson, S. Searle, Y. Zhou, F. Kokocinski, L. McDonald, R. Evans, K. Phillips, A. Atkinson, R. Cooper, C. Jones, R. E. Hall, T. D. Andrews, C. Lloyd, R. Ainscough, J. P. Almeida, K. D. Ambrose, F. Anderson, R. W. Andrew, R. I. S. Ashwell, K. Aubin, A. K. Babbage, C. L. Bagguley, J. Bailey, R. Banerjee¹, H. Beasley, G. Bethel, C. P. Bird, S. Bray-Allen, J. Y. Brown, A. J. Brown, S. P. Bryant¹, D. Buckley, D. C. Burford¹, W. D. H. Burrill¹, J. Burton, J. Bye, C. Carder, J. C. Chapman, S. Y. Clark, G. Clarke, C. Clee, S. M. Clegg¹, V. Copley, R. E. Collier, N. Corby, G. J. Coville, J. Davies, R. Deadman, P. Dhami¹, O. Dovey¹, M. Dunn, M. Earthrowl, A. G. Ellington, H. Errington, L. M. Faulkner¹, A. Frankish, J. Frankland, L. French, P. Garner, J. Garnett, L. Gay, M. R. J. Ghorri, R. Gibson, L. M. Gilby, W. Gillett, R. J. Glithero, D. V. Grafham, S. M. Gribble¹, C. Griffiths, S. Griffiths-Jones, R. Grocock, S. Hammond, E. S. I. Harrison, E. Hart, E. Haugen, P. D. Heath, S. Holmes, K. Holt, P. J. Howden, A. R. Hunt, S. E. Hunt, G. Hunter, J. Isherwood, R. James, C. Johnson, D. Johnson, A. Joy, M. Kay, J. K. Kershaw, M. Kibukawa, A. M. Kimberley, A. King, A. J. Knights, H. Lad, G. Laird, C. F. Langford¹, S. Lawlor, D. A. Leongamornlert, D. M. Lloyd, J. Loveland, J. Lovell, M. J. Lush, R. Lyne, S. Martin, M. Mashreghi-Mohammadi, L. Matthews, N. S. W. Matthews, S. McLaren, S. Milne, S. Mistry, M. J. F. M. oore, T. Nickerson, C. N. O'Dell, K. Oliver, A. Palmeiri, S. A. Palmer, R. D. Pandian¹, A. Parker, D. Patel, A. V. Pearce, A. I. Peck, S. Pelan, K. Phelps, B. J. Phillimore, R. Plumb, K. M. Porter¹, E. Prigmore¹, J. Rajan, C. Raymond, G. Rouse, C. Saenphimmachak, H. K. Sehra, E. Sheridan, R. Shownkeen, S. Sims, C. D. Skuce, M. Smith, C. Steward, S. Subramanian, N. Sycamore, A. Tracey, A. Tromans, Z. Van Helmond, M. Wall J. M. Wallis, S. White, S. L. Whitehead, J. E. Wilkinson, D. L. Willey, H. Williams, L. Wilming, P. W. Wray, Z. Wu, A. Coulson, M. Vaudin, J. E. Sulston, R. Durbin, T. Hubbard, R. Wooster, I. Dunham, N. P. Carter, G. McVean, M. T. Ross, J. Harrow, M. V. Olson, S. Beck, J. Rogers & D. R. Bentley

¹The Wellcome Trust Sanger Institute, The Wellcome Trust Genome Campus, Hinxton, Cambridgeshire CB10 1SA, UK.

Nature **441**, 315–321 (2006).

We inadvertently omitted the names of the following authors: R. Banerjee, S. P. Bryant, D. C. Burford, W. D. H. Burrill, S. M. Clegg, P. Dhami, O. Dovey, L. M. Faulkner, S. M. Gribble, C. F. Langford, R. D. Pandian, K. M. Porter and E. Prigmore.

RETRACTION

doi:10.1038/nature05298

Origin of the obliquities of the giant planets in mutual interactions in the early Solar System

Adrián Brunini

Nature **440**, 1163–1165 (2006)

When a new, independent code is used for the calculations on which the conclusions of this Letter were based, the results reported for the evolution of obliquity cannot be reproduced. This code was written in the inertial frame and is more reliable than the one used in the Letter. In most runs, the obliquities can change by only a few degrees and attain large values in only a very few cases. In addition, the obliquity variation shown in the Supplementary Information, although correct, originates from changes in the orbital inclination of the planet, and close encounters are not effective in causing large obliquities.

CORRIGENDUM

doi:10.1038/nature05296

Complete photo-fragmentation of the deuterium molecule

T. Weber, A. O. Czasch, O. Jagutzki, A. K. Müller, V. Mergel, A. Kheifets, E. Rotenberg, G. Meigs, M. H. Prior, S. Daveau, A. Landers, C. L. Cocke, T. Osipov, R. Díez Muiño, H. Schmidt-Böcking & R. Dörner

Nature 431, 437–440 (2004)

In Figs 2 and 3 of this Letter, we calculated the angle between the molecular axis and the in-plane electron as $\varphi_{e,mol} = \varphi_{mol} - \varphi_e$, where φ_e and φ_{mol} are the angles of the electron and the molecular axis with respect to the polarization vector ϵ . This distribution was then rotated by the angle of the molecular axis to picture the electron emission in the polar plots. Here a sign error occurred: instead of calculating and plotting $\varphi_{mol} - \varphi_{e,mol} = \varphi_e$, we presented $\varphi_{mol} + \varphi_{e,mol} = 2\varphi_{mol} - \varphi_e$.

Figure 1 illustrates the corrected results (compare with Figs 2, 3 in our Letter). The data now show an even stronger dependence of the angular distribution of the electrons on the molecular orientation, emphasizing the importance of the new internal reference axis. Thus, the angular distribution in Fig. 1a does not resemble a helium-like dipole pattern any more (as we stated originally), which would be aligned mainly along the polarization axis, but shows a preferred emission of electrons rectangular to the molecular axis.

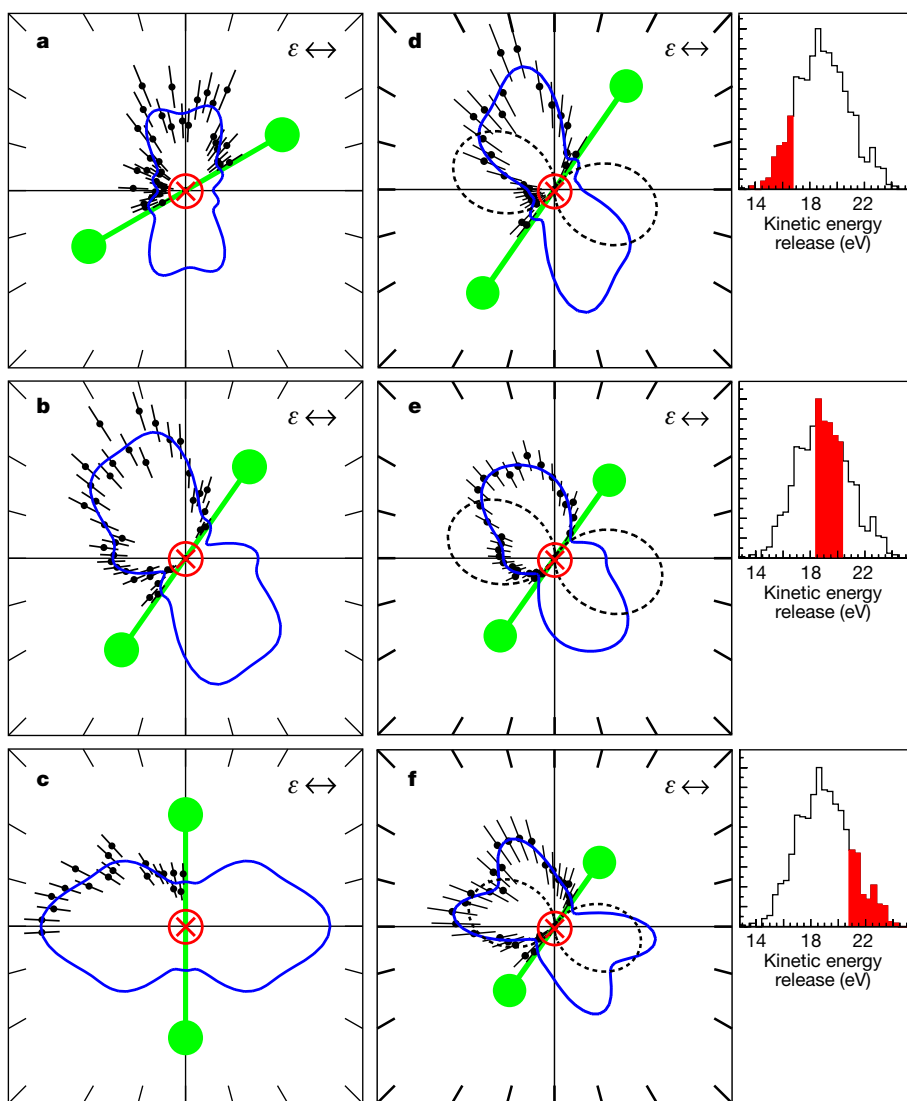


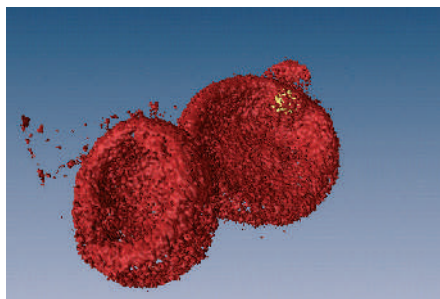
Figure 1 | Angular distribution in the non-coplanar frame as a function of molecular orientation for the photo double ionization of deuterium molecules. a–f, The angular distribution of one electron (black dots, with error bars indicating the standard deviation of the mean value) in the plane of the molecular axis (green barbell) and the electric field vector of the linear polarized light ϵ (horizontal double arrow) for a photon energy of 75.5 eV. (In **a**, the angle between the molecule and the polarization axis is 30°, whereas it is 55° in **b** and 90° in **c**.) The other electron moves orthogonally out of the plane towards the observer (red circled cross). In **d–f**, the molecular axis is fixed at 55° while the kinetic energy release varies (rightmost panels). Each electron has 12.25 eV energy. Blue solid lines show a fit with spherical harmonics ($l \in [1,4]$, $m \in [0,1]$). The dashed black lines represent the result from a single centre expansion of the molecular ground state and a CCC expansion of the final two-electron continuum.

Something to see

Light microscopy is undergoing a renaissance, with a huge range of tools and techniques for gathering biological data with unprecedented speed and resolution. Michael Eisenstein takes a closer look.

Confocal microscopy is now a well-established technique for the three-dimensional imaging of cellular structures. But despite its success, the technique has its limitations when imaging live cells. The scanning process can greatly reduce imaging speed, for example, and the powerful lasers involved can damage the cells. Some manufacturers have addressed these problems by designing alternative systems such as 'restoration' microscopy (see 'Achieving clarity', below), whereas others have strived to improve confocal technology. Leica Microsystems of Wetzlar, Germany, for instance, still uses standard point-rastering in its high-end TCS SP5 confocal instrument, but it incorporates a resonant scanner for real-time 'true confocal' imaging with little cell damage.

An alternative approach uses a rapidly rotating disk, called a Nipkow spinning disk, which has numerous apertures to illuminate hundreds of spots simultaneously. This allows faster imaging with reduced photobleaching, although it does suffer from some loss in resolution. Perkin-Elmer Life and Analytical Sciences of Boston, Massachusetts, was among the first companies to develop instruments using this technology.



High-resolution image of healthy (left) and malarial red blood cells taken with Leica's TCS-4Pi.

Its UltraVIEW ERS microscope uses parts from leading Nipkow-disk manufacturer Yokogawa in Tokyo, Japan. The system is designed for live-cell confocal imaging and can be fitted with a sensitive electron-multiplying CCD (EMCCD) camera for detection. "This offers frame-rates in excess of 100 frames per second, if you're looking at relatively small areas in your sample," says product leader Paul Orange. Olympus of Tokyo also makes a spinning-disk system, called the DSU, which features disks with slits instead of holes. Five different disks are available that vary

in terms of slit number and spacing. "You can match the slit spacing to the numerical aperture of your objective," says product manager Nicolas George. The DSU can also incorporate high-resolution EMCCD cameras for imaging live specimens at up to 150 frames per second.

The LiveScan SFC from Nikon Instruments in Melville, New York, uses arrays of pinholes or slits for multipoint imaging. These remain stationary while mirrors sweep the beam spots over the sample — a process known as swept-field confocal imaging, developed by Prairie Technologies of Madison, Wisconsin. Nikon is also introducing controlled light emission microscopy, which uses feedback from the detection process to modulate laser intensity to prevent oversaturation or unnecessary illumination of signal-free regions. "You sacrifice a little bit of temporal speed," says Stan Schwartz, vice-president of Nikon's microscopy division, "but the casual user can make correct and beautiful images, and you significantly reduce photobleaching and phototoxicity."

The LSM 5 LIVE from Carl Zeiss Micro-Imaging in Jena, Germany, uses a line-scanning approach that can image an area 512 × 512

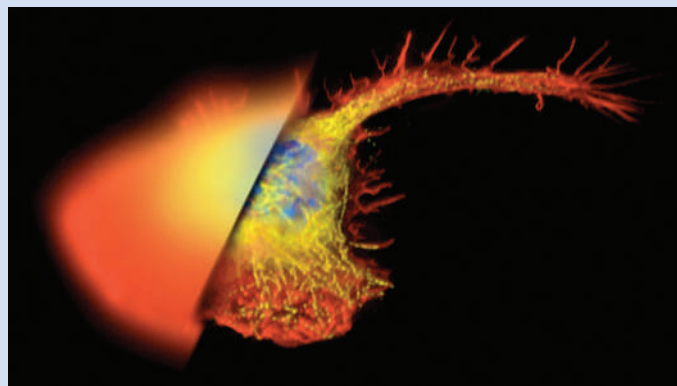
LEICA MICROSYSTEMS

ACHIEVING CLARITY

A key strength of confocal imaging is the elimination of emitted light from outside the focal point, such as that from tissue autofluorescence. But there are situations when non-confocal imaging is a better option. Deconvolution or restoration microscopy, in which computational algorithms remove artefactual blurring from actual image fluorescence while restoring out-of-focus fluorescence to its proper location, is one solution.

Applied Precision of Issaquah, Washington, was the first company to offer a complete real-time restoration-microscopy instrument, the DeltaVision RT. This combines a motorized stage with deconvolution software to collect and process two-dimensional image sections for the real-time assembly of 'restored' three-dimensional image projections.

Although restoration microscopy is sometimes presented as



Huygens software from Scientific Volume Imaging restored the detail to this macrophage imaged by wide-field microscopy.

an economical alternative to confocal, many confocal users find advantages in computational image correction, including compensation for potential resolution loss with spinning-disk instruments. Confocal manufacturers have responded by incorporating deconvolution into their software packages. For

example, Nikon Instruments in Melville, New York, offers a '2D-RT decon' module that removes blurring from two-dimensional sections to clean up three-dimensional confocal images in real time, and Olympus of Tokyo uses deconvolution tools developed by Intelligent Imaging Innovations in Denver, Colorado.

Many users also opt for dedicated deconvolution software, such as the Huygens suite from Scientific Volume Imaging of Hilversum, the Netherlands. "Our software offers as much knowledge as we dare to put in about microscope image formation and noise characteristics," explains founder Hans van der Voort, "and with that we can recover as much as possible about the original object."

Huygens is regularly updated to tackle the data being generated by new imaging techniques — an onerous task for an expanding field. Above all, the challenge is keeping the final image clean and true-to-life. "What everybody hates is a restored image that is an artefact," says van der Voort. "You can make an image like a photograph, which is nice to see. But at second glance, if you really want to analyse your data, what you want is reliability."

M.E.

J. EVANS

pixels at up to 120 frames per second, or even faster for smaller sections, which means that rapid physiological events can be visualized. According to Bernhard Zimmermann, head of product management at Zeiss, this shift from point- to line-scanning has minimal impact on resolution for most studies. "If you're looking at intracellular vesicles, you will hardly see a difference," he says.

Adventures in the *n*th dimension

These days a confocal microscope has to do more than simply track fluorophores. Multi-dimensional imaging is microscopy's new catchphrase, and the latest instruments gather information that goes beyond basic spatial orientation to provide time-lapse images and detailed fluorescence profiles.

The expansion in the range of available fluorescent labels means far more biological information can be colour-coded, and manufacturers have responded with advanced strategies for enhancing spectral resolution. For example, the LSM 510 META from Zeiss uses a 32-channel photomultiplier array behind an optical grating to generate full spectral signatures for each pixel; this is followed by further software-based separation. "You can even separate fluorescent emissions that peak at the same wavelength," says Zimmermann, "as long as the individual emission curves are different." Nikon offers a similar strategy with its C1si confocal instrument, which can achieve spectral resolution as low as two nanometres.

Leica enhances confocal spectral precision

with proprietary technologies that include the acousto-optical beam-splitter, which controls excitation illumination in a filter-free manner, and SP module, which uses a prism-based approach for precise spectral separation. These will be enhanced by its forthcoming introduction of white-light lasers, which provide coherent, continuous-spectrum illumination. "This will give us total freedom in excitation and detection for all available dyes," says Frank Olschewski, Leica's manager for confocal product development.

Fluorescence lifetime imaging (FLIM), which involves measuring the emission decay rate for an excited fluorophore, is also gaining interest for imaging. It can, for example, improve the quality of fluorescence resonance energy transfer (FRET) experiments. "When there is a FRET interaction, the donor lifetime gets shortened — this is an absolute, and it doesn't depend on the quantity or intensity," says John White of the Laboratory for Optical and Computational Instrumentation (LOCI) at the University of Wisconsin at Madison. "If you're doing other types of FRET measurements there are other reasons why you might get reduction in intensity or increased intensity, such as differential bleaching or compartmentalization." FLIM requires extremely high-speed photon detectors; Becker & Hickl of Berlin were an early leader in this regard, and Zeiss uses these detectors in its instruments. Leica also offers a FLIM attachment for its confocal instruments, which takes advantage of the SP detection system to provide 'spectral FLIM',

with precise wavelength selection.

Standard fluorescence microscopes produce a focal spot that is ovoid rather than spherical. So although high-end instruments can produce images with *x-y* resolution of up to 180 nanometres, the resolution is considerably poorer along the *z* axis — around 500 to 800 nanometres — which reduces the quality of three-dimensional reconstructions.

One way round this uses two objective lenses, where the interference between the two optical wavefronts produces an effectively spherical focal spot. Stefan Hell and his colleagues at the Max Planck Institute for Biophysical Chemistry in Göttingen, Germany, demonstrated the effectiveness of this approach with their confocal 4Pi instrument, which can achieve a resolution of 100 nm in all directions. The original 4Pi design was relatively slow, barring its use in live specimens, but Hell and his team have since developed a multifocal, multiphoton version — MMM-4Pi — that scans samples with 64 foci, and images at a rate that makes high-resolution three-dimensional imaging of live cells possible. A 'user-friendly' commercial version of Hell's 4Pi system, the TCS 4Pi, is now available from Leica.

A related method, I5M, was developed by Mats Gustafsson of the University of California, San Francisco. I5M also achieves axial resolution below 100 nm, but is based on a wide-field configuration, with interference taking place over the entire field of view rather than at a point, and can be faster and potentially brighter than confocal 4Pi. Interference imaging produces

THINKING BIG, SEEING SMALL

Most optical imaging techniques are subject to the tyranny of the diffraction limit, where the optical properties of conventional objective lenses make it impossible to distinguish two objects separated by less than 180 nanometres in the focal plane.

But limits were made to be broken. In 1993, Stefan Hell of the Max Planck Institute for Biophysical Chemistry in Göttingen, Germany, developed the concept of stimulated emission depletion (STED). Fluorescent molecules are activated with a laser spot, which is then overlaid with a ring-shaped beam of low-energy photons that shrinks the effective area of excitation to an extent determined by the brightness of the ring. This means that STED can generate almost arbitrarily small excitation areas.

The instrument lives up to its promise, achieving resolution of well under 100 nanometres along

all axes, and representing the first 'super resolution' technique to break the diffraction limit. This platform has evolved rapidly — Leica Microsystems of Wetzlar, Germany, plans to release a

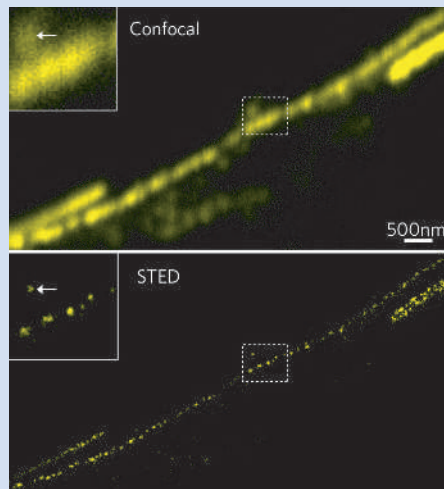
commercial instrument in 2007 — and Hell has also integrated STED with his confocal 4Pi instrument to push the resolution limit further. "We've got a resolution of 15 to 30 nanometres in the focal plane," he

says. Proof-of-concept studies suggest that STED should be suitable for live-cell imaging, although imaging rapid events with small focal volumes could prove tricky.

Fluorophore concentration is a major concern in super-resolution imaging. Mats Gustafsson, of the University of California, San Francisco, has demonstrated the use of patterned light in a wide-field

'structured illumination' scheme that is effective for planar imaging at less than 50-nm resolution, but he feels that such techniques may be less ideal for smaller focal regions. "As your resolution volume shrinks, it holds a smaller and smaller number of molecules," he says, "and the individual stochastic response of each molecule becomes more apparent. Eventually, you reach a point where it is better to exploit, rather than fight, this independent molecular behaviour. This is precisely what Eric Betzig, for example, has done beautifully."

Betzig, of the Howard Hughes Medical Institute (HHMI) Janelia Farms campus in Virginia, cites the research of Hell and Gustafsson as fuelling his interest in super-resolution. His efforts recently yielded a technique for single-molecule imaging: photoactivated localization microscopy (PALM), which he developed in



Stimulated emission depletion provides improved resolution for immunostained neurofilaments compared with confocal microscopy.

R. MEDDA, G. DONNERT, S. HELL

artefactual 'lobes', which must be removed by deconvolution; these tend to be more prominent with I5M, and lobe removal is made easier by imaging with lenses that have a high numerical aperture. These are typically oil-immersion, and use of I5M is presently restricted to fixed specimens. Gustafsson's group has also developed a variant technique, I5S, which incorporates structured illumination to surpass the 'diffraction limit' in the image plane. Together, techniques such as I5S and confocal 4Pi represent important early steps in 'super-resolution' microscopy, an increasingly vibrant area of research and technological development (see 'Thinking big, seeing small').

Nonlinear thinking

The real rising star of imaging is multiphoton microscopy (MPM), a nonlinear imaging method. In MPM, target molecules are stimulated by a pair (or more) of low-energy photons virtually simultaneously, providing sufficient energy in combination to induce excitation and photon emission. This approach uses long-wavelength, near-infrared pulsed lasers that can penetrate deep into tissues with heavy scattering properties and that are less likely to produce phototoxic side effects. Additionally, multiphoton excitation results in dramatically reduced background fluorescence. Price remains a major obstacle to adoption, as pulsed lasers are extremely expensive, but those who have tried the method generally walk away impressed. "I think it's the best technique in town for look-

ing deep into solid tissue and figuring out the dynamics of what's going on," says White.

Zeiss currently controls the patent for femto-second pulsed-laser MPM, and has made this technology available with its LSM 510 NLO. Several companies have also sublicensed this technology, and Olympus has made such an arrangement for its FV1000-MPE instrument, which was launched earlier this month. Leica also offers an MPM configuration for the TCS SP5, based on a separate patent that makes use of picosecond-pulse lasers.

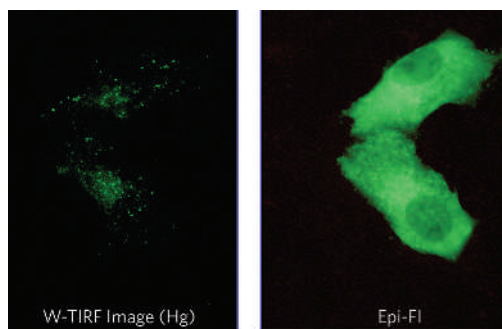
Current MPM systems offer reasonable speed for live-cell work, but LaVision's TriMScope unit further improves imaging frame-rates with a multifocal approach that simultaneously scans samples with up to 64 foci, enabling the imaging of a 1004×1002 -pixel field at 31 frames per second. TriMScope can also be adapted to perform FLIM or

second-harmonic generation imaging, another nonlinear method gaining interest for its label-free imaging capabilities.

Scratching the surface

Some of the most interesting biological events take place near the cell surface, and for imaging beauty that goes only membrane deep, total internal reflection fluorescence (TIRF) microscopy is a strong option. A sample on a glass coverslip is illuminated either through a prism or an objective lens of high numerical aperture. When light enters the coverslip at an angle greater than the critical angle, as determined by the difference between the refractive indices of the glass and the specimen, the incident beam does not enter the specimen, but instead produces an evanescent electromagnetic field. This field can excite fluorophores close to the surface (within about 200 nm) of the coverslip, allowing non-destructive, live-cell imaging at the cell membrane with single-molecule resolution.

Most major microscope manufacturers now offer multi-wavelength TIRF modules for their inverted microscopes. These typically benefit from precise computer control of the angle and position of the excitation beam. Zeiss offers a compact, stand-alone option with its Laser TIRF system. TIRF instruments typically use oil-immersion lenses with a numerical aperture of 1.45, an effective minimum for good imaging. Recently, Nikon broke this barrier with a set of Plan Apo 60x

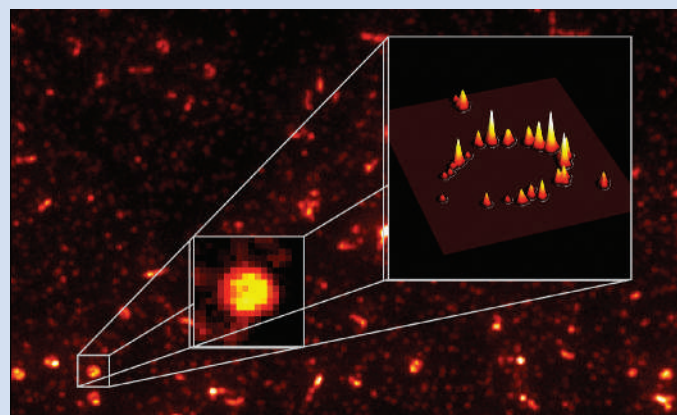


Nikon's white-light TIRF can be used to image individual clathrin-coated vesicles in the process of exocytosis.

D. AXELROD

► collaboration with Harald Hess of NuQuest Research in La Jolla, California. Subsets of individual photoactivatable fluorescent molecules are activated, and subsequently bleached, within a sample through laser exposure, and then imaged by total internal reflection fluorescence (TIRF) microscopy. This process is repeated, and the resulting images are superimposed into a stack that is computationally resolved into a fluorescent molecule 'map' with resolution below 10 nm.

The technique is not intended for live-cell work, but could offer a powerful super-resolution imaging tool for fixed cells. "There's good hope that it will be fairly easy to adapt for commercial TIRF," says Betzig. "I have some more grandiose ideas about how to do three-dimensional PALM, but the first order of business is to get a more usable instrument for cell biologists."



Imaging with stochastic optical reconstruction microscopy (inset) can reveal the detailed structure of RecA-DNA filaments.

HHMI investigator Xiaowei Zhuang, at Harvard University in Cambridge, Massachusetts, working with students Mark Bates and Michael Rust, recently described a similar technique, stochastic optical reconstruction microscopy (STORM). STORM emerged from the discovery

by Zhuang's group that the commonly used Cy5 fluorophore has photoswitchable properties. "It can be switched in a controlled way, back and forth hundreds of times, between light and dark states," says Zhuang. Like PALM, STORM involves cycles of selective excitation and imaging

followed by computational reconstruction of the full image. As the fluorophores can be turned on and off, it is possible to do time-resolved imaging and faster imaging cycles. Resolution as low as 18 nanometres has been demonstrated, and Zhuang thinks this is just the beginning. Right now, her top priorities are exploring the potential for multicolour imaging, and improving imaging speed. "If we can bring STORM to one-second resolution, that's going to open up a big door and enable many different things," she says.

Surveying the field, Hell views these various approaches as complementary tools that advance a common goal — forcing researchers to re-examine what is possible in imaging. "I think breaking the diffraction barrier is a fundamental step forward," he says. "This is an idea whose time has been coming for a long time now."

M.E.

X. ZHUANG

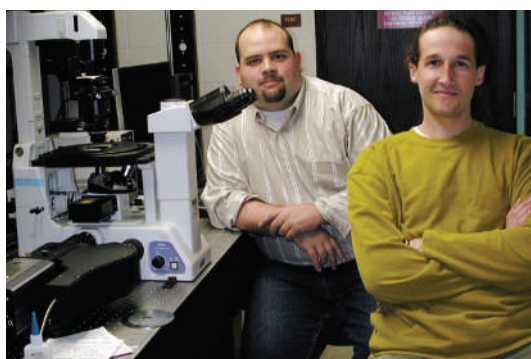
and 100× magnification objectives. “We achieved a numerical aperture of 1.49,” says Schwartz, “and that’s just under the theoretical limit for imaging with glass.”

Putting it all together

Modern imaging experiments have become very complex, generating huge amounts of multidimensional data. Although microscope manufacturers will typically provide reasonably effective tools for image processing and analysis, many users also opt for third-party software with more extensive analytical capabilities.

MetaMorph from Molecular Devices of Sunnyvale, California, integrates hardware control with a large toolbox of resources for multidimensional imaging, including a range of ‘application modules’ for image processing. “These are specific segmentation and analysis modules where, rather than having the user write macros, we’ve automated the entire process by having one dialogue box that helps segment your image and gives you relevant measurements,” says product manager Magali Tranié.

Imaris from Bitplane in Zurich, Switzerland, also uses a module-based approach for segmentation and analysis of multidimensional images. “You can click on any object that you see and immediately, in the same screen, get the statistics for that particular object,” says vice-president and director of sales Michael Wussow. “The same is true in reverse — we have an interactive sorting tab, where you move a bar on a histogram and things will appear or



Kevin Eliceiri and Curtis Rueden have worked together to develop open-source solutions for the imaging community.

disappear depending on whether or not they meet your criteria.” Bitplane also allows users to code their own routines, and hosts a number of user-generated plug-ins on its webpage.

Improvision of Coventry, UK, recently released the latest version of its Volocity package, which consists of four products — Acquisition, Visualization, Quantitation and Restoration — that can also be integrated. Among Volocity’s strengths are its object detection and tracking capabilities, and powerful rendering techniques. “Every voxel within the image set is computed by the Volocity rendering engine,” says senior marketing specialist Nicky Francis. “So it isn’t a surface rendering technique, but a fully interactive volume-rendering technique. This makes it much more true-to-life.”

Numerous other powerful commercial options are available, but the academic com-

munity has also stepped up to provide a variety of open-source solutions. Among the most popular is ImageJ, developed at the US National Institutes of Health by Wayne Rasband. ImageJ is a multipurpose imaging tool, maintained by Rasband but powered by a highly active user community. “There are about 400 plug-ins on our webpage, contributed by more than 100 people,” says Rasband. Another popular program is VisBio, developed by the LOCI’s Curtis Rueden for working with multidimensional data. According to Kevin Eliceiri, co-director of the LOCI, such tools complement commercial products. “The goal is to fill in the gaps,” he says. “There are needs that often can-

not yet be met by the commercial community because they either represent too small a market or are emerging techniques.” Such projects have also meshed with the efforts of the Open Microscopy Environment (OME) to develop tools that aid collection and sharing of complex image data (see ‘Tower of Babel’).

Software — and hardware — requirements will only grow more severe as scientists attempt to coax increasingly complex data from smaller numbers of photons. Technology aside, the underlying challenges for this field remain clear and simple. “There are three issues in microscopy that never change — I want the best resolution, I want the best sensitivity and I want the best speed,” says Nikon’s Schwartz. “And it’s really difficult to get all three.” ■

Michael Eisenstein is technology editor for *Nature* and *Nature Methods*.

K. ELICEIRI

TOWER OF BABEL

The Open Microscopy Environment (OME) first emerged from the recognition that an explosion in imaging data was imminent. Image files were becoming bigger and more complex, representing copious data on numerous parameters as well as experiment-specific ‘metadata’. “The ability to provide links between an image, any processed versions of an image, the data describing the acquisition of that image, and any analytical results generated about that image is a critical capability,” says Jason Swedlow of the University of Dundee, UK.

When Swedlow, along with Ilya Goldberg at the US National Institute on Aging, Peter Sorger at the Massachusetts Institute of Technology, and researchers at the Laboratory for Optical and Computational Instrumentation (LOCI) at the University of Wisconsin at Madison, officially

launched the OME in 2001, the development of a universal file format was a top priority. “Proprietary file formats are one of the biggest problems in modern microscopy,” says the LOCI’s Kevin Eliceiri. “There are about 30 or 40 major microscopy file formats, so we’re speaking all of these different languages — it’s the Tower of Babel.”

To combat this, the OME has developed a file format called OME-XML, which retains both image pixel data and experimental metadata in a readable XML-based file. The LOCI has since refined this format as the OME-TIFF, which still encapsulates metadata in XML but stores pixel data in a TIFF format. “We have the best of both worlds now, in that TIFF is probably the closest there is to a universal image format,” says Eliceiri.

The OME also offers open-source software for image management and analysis, which



Jason Swedlow helped launch the Open Microscopy Environment.

uses these new formats. Swedlow acknowledges that OME software may prove challenging for less computer-savvy users, but says the developers are working hard to make it easier to use. Meanwhile, a growing number of commercial packages now not only recognize the OME formats, but also record data into them. Indeed, the efforts of Swedlow and his colleagues

have won accolades from several manufacturers. “Maybe up to 15 or 20% of our development resources go towards reader updates,” says Michael Wussow of Bitplane in Zurich, “so we really like this idea of an Open Microscopy Environment.”

The OME also intends to tackle another weighty issue: handling and browsing increasingly bulky data sets. “We currently store about 50 terabytes of data from our imaging facility,” says Swedlow, “and the OME’s goal is to provide software to handle data on this scale.” Of course, software isn’t the whole answer, and significant investment in hardware and network infrastructure will be necessary for labs serious about imaging. “An enterprise-scale data-storage facility doesn’t sound like hypothesis-driven research,” says Swedlow, “but it is a required tool for hypothesis-driven research using imaging.” ■

J. SWEDLOW

M.E.

naturejobs

**THE CAREERS
MAGAZINE FOR
SCIENTISTS**

There's both good news and bad news on the UK academic salary front. On the plus side, academic salaries across Britain average £36,500 (US\$68,700), in line with faculty wages in the United States and Western Europe, at least on the entry level, according to an analysis of the UK's Higher Education Statistics Agency staff record for 2004–05 published by *The Times Higher Education Supplement*.

On the minus side, there's a difference of some £21,000 between average salaries at the top and bottom of the table (and that's without putting the £114,000 average salaries at the London Business School into the equation).

The apparently wide variations between institutions must be treated with caution, however. Four of the six highest-paying universities — with average salaries of over £40,000 — are based in London. That reflects the higher cost of living there, as well as a drive by Imperial College and University College London to be more internationally competitive. The lowest payer, Lancaster University, with an average salary of £21,686, is based in the north, but Lancaster's administrators say that they included pay for postgraduate researchers. Other variations can be at least partially accounted for by universities that employ large numbers of research staff, such as the University of Oxford (£33,879) and Durham University (£35,201).

The real bad news is still for female academics. Even those of professorial rank are being paid 6% less than their male counterparts, and some institutions are paying women 10% less than men.

Despite the caveats, the analysis provides an impetus for all UK universities to re-examine their pay structure and ask themselves whether young researchers and women are fairly compensated. Universities with average salaries near the national average might congratulate themselves — but they should look at the range of the salaries they offer. And to be even more competitive in attracting researchers, they could use an international benchmark, rather than a national one. Doing so could make Britain more desirable for researchers weighing global career options.

Paul Smaglik, *Naturejobs* editor

Publisher: Ben Crowe
Editor: Paul Smaglik
Assistant Editor: Gene Russo

European Head Office, London
The Macmillan Building,
4 Crinan Street, London N1 9XW, UK
Tel: +44 (0) 20 7843 4961
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

European Sales Manager:
Andy Douglas (4975)
e-mail: a.douglas@nature.com
Business Development Manager:
Amelie Pequignot (4974)
e-mail: a.pequignot@nature.com
Natureevents:
Claudia Paulsen Young (+44 (0) 20 7014 4015)
e-mail: c.paulsenyoung@nature.com
France/Switzerland/Belgium:
Muriel Lestranguez (4994)

UK/Ireland/Italy/RoW:
Loredana Milanese (4944)
Nils Moeller (4953)
Scandinavia/Spain/Portugal:
Evelina Rubio-Morgan (4973)
Germany/Austria/The Netherlands:
Reya Silao (4970)
Online Job Postings:
Matthew Ward (+44 (0) 20 7014 4059)

European Satellite Office
Germany: Patrick Phelan
Tel: +49 89 54 90 57 11
Fax: +49 89 54 90 57 20
e-mail: p.phelan@nature.com

Advertising Production Manager:
Stephen Russell
To send materials use London address above.
Tel: +44 (0) 20 7843 4816
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

Naturejobs web development: Tom Hancock
Naturejobs online production:
Catherine Alexander

US Head Office, New York
75 Varick Street, 9th Floor,
New York, NY 10013-1917
Tel: +1 800 989 7718
Fax: +1 800 989 7103
e-mail: naturejobs@natureny.com

US Sales Manager: Peter Bless

Japan Head Office, Tokyo
Chiyoda Building, 2-37 Ichigayatamachi,
Shinjuku-ku, Tokyo 162-0843
Tel: +81 3 3267 8751
Fax: +81 3 3267 8746

Asia-Pacific Sales Manager:
Ayako Watanabe
e-mail: a.watanabe@natureasia.com



NORTHERN EXPOSURE

An alliance of universities in the north of England hopes to transform Britain's former industrial heartland into a centre of scientific excellence. **Paul Smaglik** reports.

The industrial north of England fuelled the British Empire with coal and steel, but an initiative funded last week aims to transform the region into a globally competitive science centre. A £6-million (US\$11-million) grant, awarded by the Northwest Regional Development Agency, will forge collaborations between eight universities in Sheffield, Manchester, Leeds, Liverpool, Newcastle, Durham, Lancaster and York. If it's successful, the alliance could create hundreds of jobs and attract millions of pounds in additional research funds.

Called the N8, the alliance intends to help these universities compete better for global scientific funding, infrastructure and industry relationships. The seed money will pay for administrative support for directors of five research areas and the creation of an independent umbrella company to develop intellectual property. It will also fund technology-transfer specialists who can package intellectual property from the members and work with companies and grant-giving agencies to pursue further funding. The N8 intends to build on research strengths in five areas — regenerative medicine, energy, water, ageing research and molecular engineering (including nanotechnology).

The N8 plan is part of the Northern Way, a 20-year strategy launched last year by the UK government to transform the economy of the north of England by moving it towards more knowledge-based industry. Administered through the three regional development agencies, the Northwest Regional Development Agency, One Northeast and Yorkshire Forward, the Northern Way's goals are lofty — to use £100 million of seed funding to close the £30-billion gap in output between north and south by 2025.

There are plenty of obstacles to be overcome before that, however. The north is in competition with the 'golden triangle' of Oxford and Cambridge, the London universities and the associated hi-tech industries in the south of England. On the political front, governmental and bureaucratic changes may slow or hinder the collaboration, while on the corporate side, the N8 institutions must establish a stronger track record in creating new high-tech businesses.

Perhaps the most immediate barrier is a history of competition and the suspicion that comes with it, say

some of the pro vice-chancellors involved. Many of the universities have their own aspirations beyond the N8. "There are probably eight different motives for being involved in the N8," says Simon Gaskell, University of Manchester pro vice-chancellor of research. "But there's nothing wrong with that."

David Secher, who has experience in technology transfer at both the Medical Research Council and the University of Cambridge, has been brought on board as chief executive to balance the interests of the eight institutions. "The key thing is that every university and research centre has to see some benefit," he says.

It's one thing to agree on common research goals, but quite another to give up intellectual property that could benefit one university in the hope that bigger returns could be realized collectively. "It's not realistic to say, 'Let's put all our intellectual property into the N8,'" Gaskell says. "But some collaborations with the N8 will be appropriate."

Intellectual ambitions

To help balance each university's interests, Secher has created a separate holding company for shared intellectual property arising out of the collaboration. Putting together intellectual-property portfolios to tempt industry may make sharing necessary, and it should create an incentive for members to cooperate on pursuing funding from business and government for research and infrastructure.

Secher admits that keeping the N8 together won't be an easy job — especially as some big hitters, such as the University of Manchester, could outshine their partners. And resentment could build if smaller players get sidelined because of others' ambitions. But the eight pro vice-chancellors, who began discussing the N8 concept informally last year, have already begun to address this problem. "The smaller universities accept that they may not have strengths in all five research areas," says Secher.

Members of the N8 have had some success in raising funds within smaller networks. A collaboration between the University of Newcastle upon Tyne and Durham University is bringing together research teams studying embryonic, cord blood, germline and somatic stem cells, says Trevor Page, Newcastle's pro vice-chancellor. It's part of a group called the North



David Secher: balancing act.



East England Stem Cell Institute, which recently netted £10 million from One Northeast for translational research. And Jon Saunders, a pro vice-chancellor at the University of Liverpool, notes that Liverpool and Lancaster have received research funds to turn waste into energy. "We were doing this independently of the N8," Saunders says. "It just happened to fit."

Joining forces may be even more important in attracting large national laboratories, which mean hundreds of jobs and millions of pounds in funding, to the north. "Unless the N8 becomes known as a centre of excellence, the big national facilities will go to the south," Secher says.

Big science

Secher and the N8 steering committee have set their sights on one project in particular: a new synchrotron light source for elucidating the structures of proteins and other materials. The north could also benefit from a shared mass-spectrometry facility. "Most northern universities would not be able to afford either of these on their own," Secher says.

The north has suffered some disappointments in this area. When the new Diamond synchrotron light source was built in the south, it angered those working at the older synchrotron at the Daresbury Laboratory in Cheshire. Many workers didn't want to relocate to the new facility because of the higher cost of living (see *Nature* 402, 451; 1999).

Nevertheless, a new centre arose out of the ashes — the National Centre in Accelerator Science at Daresbury. "That was on the basis of three universities putting their resources together," says Saunders — in this case Liverpool, Manchester and the University of Lancaster. Other regional efforts have already set up new facilities in the north, including the National Biomanufacturing Centre, a bioprocessing centre that makes proteins for industrial research, in Speke, near Liverpool.

Some universities have invested heavily in new facilities, which could put them in competition with each other in the future. Manchester has put £650 million over the past five years into a new photonics facility, an interdisciplinary research centre, a new nursing building and renovations to a physics building. And both Manchester and Newcastle have

The University of Manchester's Incubator Building for biotech start-ups (above) and (inset) the Northern Institute for Cancer Research at Newcastle.



been awarded government funds for separate systems-biology institutes.

The N8 research centres will become self-sustaining financially, by winning national and international competitive funding from public and private sources, says Secher. He estimates that each of its five research areas could reach some £30 million a year in the coming decade.

On a national level the picture is uncertain; with Prime Minister Tony Blair due to step down next year, it is unclear how this will affect political support. If, as many predict, Gordon Brown, the present Chancellor of the Exchequer and one of the architects of the N8 initiative, takes the reins, the project is likely to receive sustained funding — if the top job goes to someone else its future is less secure.

But even if government provides extra funding, there are concerns about how quickly the money will be allocated. "For the past 15 months we've been caught in a holding pattern while the three regional development agencies in the north have established funding mechanisms," says Jones. "It's been really frustrating."

Jones also worries that scientists, who prefer to initiate their own work, may resist having their research directions and priorities imposed from above. "It's always difficult to move down from the top with a major project that the scientific community may not see an immediate benefit from," he says.

At present, companies founded on university research are far less common in the north than the south. One potential success is the University of Manchester's spin-off Renovo, which specializes in wound healing.

The real challenge, says Chris Henshaw, pro vice-chancellor of research at York University, is to convince the international business community to recognize not just the N8's potential, but its present accomplishments, and to persuade them to invest locally and set up companies in the area.

"The dream scenario is that in two years' time we will get not only UK investment, but foreign and business investment," says Henshaw. "If we can't succeed in that, then we will have to pause and rethink."

Paul Smaglik is *Naturejobs* editor.



Sustainable architecture: the award-winning Devonshire Building at the University of Newcastle.

N. HIGHAM/ALAMY; ROBINSON LIBRARY, NEWCASTLE UNIV.

ROBINSON LIBRARY, NEWCASTLE UNIV.

ADVERTISEMENT FEATURE



Lancaster University was rated in the top ten research institutions in the most recent RAE and leads pioneering work in science, technology and management on a global scale.

Lancaster has invested £200m into its estate and centres of excellence around environmental science, leadership, ICT and business outreach.

Recent high profile appointments around health, medicine and leadership have been made and the University is currently recruiting senior academics in the areas of bio-medicine, management and design.

Lancaster University is pleased to be working with its N8 partners in the delivery of the 5 initial themes.

Sustainable Water Use

Lancaster is co-ordinating the theme for sustainable water use and is home to the Centre for Sustainable Water Management (CSWM) based in the Lancaster Environment Centre, a joint venture between the Natural Environment Research Council's Centre for Ecology and Hydrology and the University. The CSWM delivers cross-disciplinary science in water-related areas to build solutions to current and future problems such as flooding, diffuse agricultural pollution, and the medium-term impact of climate change on water resources and water quality. This sits alongside other centres in sustainable energy, agriculture, environmental informatics and chemical management.

Micro-engineering

Our strengths in nano-science and nano-technology at Lancaster allow us to design, develop and engineer materials down to the atomic level.

Lancaster scientists from Physics, Engineering and Biology disciplines lead many cutting edge international collaborations involving numerous universities, research institutes and industrial partners world-wide.

Research projects include: the study of nano-science to detect abnormal cells in cancer and other diseases, the development of future nanoscale electronic devices, the behaviour of superfluids at ultra-low temperatures, development of high resolution scanning probe microscopy techniques, fabrication of novel semiconductor lasers and photodetectors for use in sensors and imaging, as well as design for micro and nano manufacturing.



Developments in Health Research and Medicine contribute to the ageing theme

The Institute for Health Research (IHR) exists to conduct high-quality health-related research and leads in areas such as public health, mental health and end of life care. It has an important role in Medical Education that has just begun on campus.

The UK's first professor of Hospice Studies is based in the International Observatory on End of Life Care and will lead on important new research to help improve the lives of the many thousands of patients and families who depend on hospice care every year.

The Observatory – part of Lancaster University's Institute for Health Research – is an International research centre



THE QUEEN'S
ANNIVERSARY PRIZES
2005

LANCASTER
UNIVERSITY



specialising in the development of hospice and palliative care.

Business, Enterprise and Innovation

Lancaster blends the traditional university focus on teaching, research and scholarship with business, enterprise and innovation and is a leader in the world's knowledge driven economy.

InfoLab21 is a world class Centre of Excellence for research, development and commercialisation of ICT.

The University is internationally respected for its work in ICT ranging from signal processing and coding, through wired and wireless computer networking including Internet, Web and multimedia technologies and ubiquitous and mobile computing, to software systems engineering and human-computer interfacing. Through InfoLab21's Knowledge Business Centre, spin-outs and start-up companies are fully supported together with an active commitment to the transfer of knowledge and technology into regional companies.



The Lancaster Leadership Centre is based in Lancaster University's world-class Management School and is the host for the Northern Way's Northern Leadership Academy, a collaboration between Lancaster, Liverpool and Leeds Universities. Its work is pioneering new ways of building leadership capability in organisations across the world.

The Academy, a key component of the Northern Way Growth Strategy, aims to build a more entrepreneurial North by providing new opportunities for leadership development for existing leaders and those with leadership potential throughout the North, in business and within communities.

Lancaster Environment Centre

A third phase of expansion to the LEC will enable a significant, measurable impact on the economic development and prosperity of the North West by providing office and laboratory facilities for the physical co-location of new ventures and established businesses and collaborative research with larger national and international organisations. This new development will focus on collaborative working with companies working in a range of sectors including: environmental technology, chemicals management, sustainable energy systems, water and wastewater management, agri-business, nuclear decommissioning and contaminated land assessment and remediation.

Science Park

Plans for development of a new Science Park adjacent to Lancaster University's Bailrigg campus are moving forward. The Science Park, of great importance to the local and regional economy, would build links to the University's major centres of expertise in information technology, environmental research and business leadership and enterprise development, and create opportunities for the University's cutting edge research, consultancy and other services to benefit businesses of all sizes.

SP90390



THE UNIVERSITY of LIVERPOOL

Department of Chemistry

Postdoctoral Researchers (4 Posts)

£27,193 - £28,849 pa

Four Postdoctoral Researcher positions are available in the EPSRC Complex Materials Discovery Portfolio Partnership. Specific areas of interest are: materials physics of new electronic and magnetic materials; crystallographic characterisation of new materials by transmission electron microscopy; synthesis and characterisation of ionic and mixed conducting materials; development of new materials for photocatalytic water splitting; synthesis of nanostructured inorganic materials; synthetic solid state chemistry.

A PhD in chemistry, materials science or physics and an excellent publication record are essential, skills in the specific areas identified above are desirable. Applications from well-qualified candidates with interests in other areas of inorganic and hybrid materials chemistry are also encouraged. The posts are available for one year initially.

Quote Ref: B/876/N

Closing Date: 17 November 2006

School of Biomedical Sciences
Department Of Physiology

Research Technician Grade D (Maternity Leave)

£17,461 - £19,073 pa

You will run a recently established Biomedical Electron Microscopy Unit (www.liv.ac.uk/emunit) and be responsible for ensuring the efficient operation of the Unit. You will be expected to train new EM users and contribute to ongoing EM based research programs. You should be educated to BTEC Higher level (or equivalent) and have familiarity with a broad range of techniques associated with cell biological EM research. Applications from graduates with a degree in a relevant subject would be welcomed. The post is to cover a period of maternity leave and is expected to last up to 15 months.

Quote Ref: A/184/N

Closing Date: 10 November 2006

For full details, or to request an application pack, visit
www.liv.ac.uk/university/jobs.html or email: jobs@liv.ac.uk
Tel 0151 794 2210 (24 hr answerphone),
please quote Ref: in all enquiries.

COMMITTED TO DIVERSITY AND EQUALITY OF OPPORTUNITY



Avoid getting
in it with
impressive
interview and
resume/CV advice

naturejobs



UNIVERSITY OF LEEDS

Faculty of Biological Sciences

Institute of Molecular and Cellular Biology

Research Fellow

Cryo-electron microscopy of recombinant dyneins

This post is available from 1 January 2007 for a fixed term of four years in Dr Stan Burgess' lab working on cryo-electron microscopy of recombinant cytoplasmic dynein motor protein, funded by BBSRC. Experience of cryo-electron microscopy and three-dimensional image processing of single particles is highly desirable. Cryo-microscopy will be performed using a Technai F20 microscope fitted with a 4k CCD camera. Some negative stain work will also be used for initial characterisation of constructs. Image processing will be performed using SPIDER, Imagic and EMAN on a 150-processor Linux cluster and numerous graphics workstations. A PhD in structural biology, electron microscopy, biochemistry, image processing or a related area, and experience with and the ability to use technically sophisticated equipment are essential.

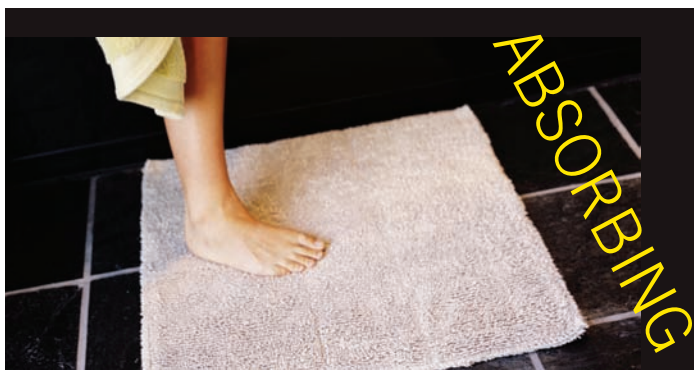
University Grade 6 (£20,842 - £24,161 p.a.) or Grade 7 (£26,402 p.a.) depending on experience

**Informal enquiries to Dr Stan Burgess tel +44 (0)113 343 4170
email s.a.burgess@leeds.ac.uk**

**To apply online please visit <http://www.leeds.ac.uk> and click on jobs.
Alternatively application packs are available from Mr A Bateman, Faculty Staff Recruitment Office tel +44 (0)113 343 8040 email fbsjobs@leeds.ac.uk
Job ref 313114 Closing date 16 November 2006**

We welcome applications from all sections of the community. Textphone for deaf applicants only 0113 343 4353. All information is available in alternative formats please contact 0113 343 4146.

WORKING TOWARDS EQUALITY AND DIVERSITY www.leeds.ac.uk



Career articles from Naturejobs

- **Naturejobs Prospect:** quick takes on career implications of current events
- **Naturejobs Special Report:** examinations of jobs issues on both sides of the bench
- **Naturejobs Careers & Recruitment:** discipline-by-discipline exploration of opportunities
- **Naturejobs Regions:** tours of scientific hubs
- **Naturejobs Movers:** traffic reports that follow high-profile scientific globe-trotters and sector-hoppers
- **Naturejobs Postdocs & Students:** guides to taking a step toward a permanent position

www.naturejobs.com

naturejobs
making science work

nature publishing group **npg**

Faculty of Life Sciences

Healing Foundation Fellowship (2 posts)

Salary £28,010 - £39,935 p.a.

Ref: LS/142/06

These two prestigious Research Fellowships are based within the newly established Healing Foundation Centre at the University of Manchester. You will have a research profile in any relevant area, such as developmental biology, healing, regeneration in any biological system, inflammation, stem cells or angiogenesis and will have the ability to demonstrate outstanding potential as an independent scientist. Applications from individuals using interdisciplinary research programmes are encouraged.

Information on The Healing Foundation can be found at www.thehealingfoundation.org

The University offers a superb environment with state-of-the-art facilities. Wound healing and tissue regeneration are important research themes for the University, as evidenced by existing (molecular, cellular, clinical) internationally renowned groups. There are particular strengths in tissue engineering, wound healing, developmental biology, cell-matrix biology and in relevant clinical sciences (e.g. plastic and reconstructive surgery). The University encourages a strong culture of interdisciplinary collaboration, thereby facilitating the rapid translation of basic science to clinical application. Co-location of all relevant groups on a single campus with new laboratory buildings and major commitments to future developments, offer an environment second to none.

Both posts are tenable for five years and come with full salary support for the fellow, together with generous funding for consumables, equipment and a staff member.

Informal enquiries may be addressed to Professor Enrique Amaya on +44 (0) 161 275 1716 or email: enrique.amaya@manchester.ac.uk

Closing date: 30 November 2006.

Postdoctoral Research Associate (2 posts)

Salary £25,633 - £28,849 p.a.

Refs: LS/153/06 & LS/154/06

Research Assistant

Ref: LS/155/06

Salary up to £23,457 p.a.

These posts are based in the laboratory of Professor Nancy Papalopulu (www.ls.manchester.ac.uk/research/themes/devbio/). The laboratory is located in a brand new building and a modern, interactive environment with state-of-the-art research facilities.

Ref: LS/153/06: You will investigate the integration of growth factor signalling and transcriptional activity in the developing telencephalon, focusing on the transcriptional and post-transcriptional regulation of the telencephalic transcription factor FoxG1. Experience in molecular biology and/or biochemical methods is required. Experience in working with *Xenopus* and imaging is an advantage but is not essential, as training will be provided. This post is funded by the Wellcome Trust and is available immediately.

Ref: LS/154/06: You will work on a fundamental question of stem cell research, the embryonic origin of neural progenitors. This project involves the development of long-term fate mapping in the *Xenopus* central nervous system with the cre/lox system. Some experience in micro-manipulation and a meticulous approach to experimentation would be an advantage. This post is funded by the BBSRC and the start date is negotiable.

For both of the above posts, you must hold a PhD in a biological discipline. A first author publication in an international journal is desirable. Please provide the names and addresses of three referees.

Ref: LS/155/06: Funded by the BBSRC to assist with the project of LS/154/06 and to provide general laboratory support. You will have some experience in molecular and cell biology to assist particularly in the generation and use of antibodies and possess excellent organisational skills.

All three posts are tenable for up to three years.

For informal enquiries, please contact Professor Nancy Papalopulu, email: nancy.papalopulu@manchester.ac.uk

Closing date: 24 November 2006.

For all the above posts, application forms and further particulars are available from our website. If necessary, you can obtain them from +44 (0) 161 275 8836 or email: Lifesciences-hr@manchester.ac.uk

Division of Epidemiology & Health Sciences

Postdoctoral Research Associate in Genetic Epidemiology

Salary £25,633 - £31,525 p.a.

Ref: MHS/419/06

The Arthritis Research Campaign's Epidemiology Unit seeks a postdoctoral research associate to contribute to our "Genetics of Musculoskeletal Disorders" research programme. You will assist in identifying genetic factors associated with a range of bone and musculoskeletal pain disorder utilising high throughput genotyping technologies and current analysis tools to examine genetic associations and gene-environment interactions of pain and bone disorders in population samples.

You will be an enthusiastic and dynamic scientist with experience in the genetics of complex diseases to take responsibility for all aspects of the programme. You will be encouraged to contribute to the leadership and development of this new area of work.

You should hold a good (2:1 or above) science degree and a PhD in genetic epidemiology or a related area.

This is for one year in the first instance and is subject to additional funding.

Informal enquiries should be directed to Dr John McBeth on +44 (0) 161 275 5788 or email: john.mcbeth@manchester.ac.uk

Application forms and further particulars are available from our website. If necessary, you can obtain them from +44 (0) 161 275 5037 or email: arcrecep@listserv.manchester.ac.uk

Closing date: 17 November 2006.

Division of Cardiovascular and Endocrine Sciences

Research Assistant/Postdoctoral Associate Position in Cellular Cardiology

Salary £21,467 - £24,886 or £25,633 - £31,525 p.a.

Ref: MHS/420/06

A post to 'map' the functional and molecular properties of the sinoatrial and atrioventricular nodes of the heart is available immediately for up to 16 months. Depending on your expertise, you will use a range of techniques including electrophysiology (optical mapping and possibly single cell patch clamp), biochemistry (Western blot and immunohistochemistry) and molecular biology (quantitative real-time and in situ hybridisation).

You will be highly motivated, have a keen interest in cellular cardiology and have a relevant PhD and/or BSc. You will have good communication skills, be able to work well both in a group and independently and have experience in one or more of the following: electrophysiology, biochemistry and molecular biology.

Informal enquiries can be made to Professor M R Boyett on +44 (0) 161 275 1192 or email: mark.boyett@manchester.ac.uk and Dr H Dobrzynski on +44 (0) 161 275 1182 or email: halina.dobrzynski@manchester.ac.uk

Application forms and further particulars are available from our website. If necessary, you can obtain them from +44 (0) 161 275 1197 or email: julie.a.heydon@manchester.ac.uk

Closing date: 10 November 2006.

For all posts, please quote the reference number.

The University will actively foster a culture of inclusion and diversity and will seek to achieve true equality of opportunity for all members of its community.

SP90446R

MOVERS

Carl Pilcher, director, NASA Astrobiology Institute, Moffett Field, California



2006: Senior scientist for astrobiology, Astrobiology programme, NASA Headquarters, Washington DC

2001-05: Senior scientist for astrobiology, Universe/Astrophysics Division, NASA Headquarters, Washington DC

Carl Pilcher proves that strategic career moves can help scientists navigate across disciplines to a dream job. Once a chemistry undergraduate, he has delved into astronomy, science policy and eventually into the nascent field of astrobiology during his 30-year career.

Once Pilcher realized his true interests lay in outer space, the first-year chemistry graduate student at the Massachusetts Institute of Technology, Cambridge, began working with a cosmochemist and a planetary astronomer, both interested in the formation of the Solar System. He used his knowledge of spectroscopy to characterize the surfaces of Jupiter's galilean satellites, and his team went on to discover frozen water in Saturn's rings and on three of Jupiter's moons, including Europa.

After 12 years at the University of Hawaii's Mauna Kea observatory, he decided to use science to contribute to public policy, and completed a master's in public affairs at Princeton University, New Jersey. Intent on moving to Washington DC, Pilcher soon became the science director of NASA's Office of Exploration Systems — a programme designed to identify possible human-based missions to the Moon or to Mars. Later he moved to NASA's space-science office where he helped to create the successful unmanned missions to Mars in the 1990s.

After claims were made about biological signatures in a Mars meteorite found in Antarctica, NASA created an astrobiology programme, which piqued Pilcher's interest. He immersed himself in a biology bootcamp including a seven-week intensive microbial-physiology programme. A few years after that he became NASA's senior astrobiologist.

Pilcher is excited to promote the field in his newest role as director of NASA's Astrobiology Institute in Moffett Field, California. The astrobiology programme faces a 50% cut in the proposed budget for this fiscal year — in part because NASA administrators are concerned about its lack of focus. Ralph Pudritz, director of McManus University's Origins Institute in Ontario, Canada, suggests that the international astrobiology scene depends on the continued success of NASA's programme. "A discovery of life on Mars could transform astrobiology from an exciting but tentative field of science into as hard a science as you could imagine," he says.

Pilcher agrees. "One of our greatest accomplishments of the twenty-first century will be understanding the biological potential of the Universe," he says. ■
Virginia Gewin

RECRUITERS & ACADEMIA

Made-to-measure postdocs

Eight months into my postdoc, I am in Berlin, diligently learning German and managing my research schedule at the Berlin Natural History Museum and the Max Delbrück Center for Molecular Medicine. Walking past the museum's collections of pickled frogs, I realize that I'm a long way from the University of Utah, and even farther from the physical-chemistry research I completed as an undergraduate.

In a quest to link my PhD chemical-biology training with my desire to be a research scientist at a natural-history museum, I designed a two-year postdoctoral project. It focused on the investigation of the peptidic toxins of venomous marine cone snails and turrids. My short-term posts include the time in Berlin, two field expeditions to Panama in association with the Smithsonian Tropical Research Institute, and a museum internship at the Paris Natural History Museum. I'm working with a global network of research scientists to expand my training in peptide chemistry and acquire knowledge in neuroscience, taxonomy and systematics.

Together with the Utah Museum of Natural History, I created a series of educational programmes and exhibition pieces to inform high-school students about the

biodiversity of cone snails and turrids, as well as the chemical biology of their toxins. I owe my scientific freedom, in part, to the postdoctoral grant funding from the US National Science Foundation's Discovery Corps Fellowship (DCF).

Tantamount to a Peace Corps for scientists, a DCF offers recent PhD candidates and mid-career scientists one- or two-year grants that support both their scientific research and a service-oriented outreach project that demonstrates different aspects of science to a larger lay audience. It allows scientists and engineers to venture beyond the lab. For example, Joseph Fortunak, a chemistry professor at Howard University, Washington, is expanding his research by working with Nigerian scientists and industry to promote 'green chemistry' education and to manufacture microcrystalline cellulose from elephant grass and other biorenewable resources.

In general, DCF awardees attempt to translate laboratory achievements into an expanded set of professional skills and a programme that serves society. Such opportunities help reinvent the humdrum of lab work and construct new routes to success. ■

Mandë Holford is a postdoctoral fellow at the University of Utah.

GRADUATE JOURNAL

How studies can save a life

What's the long-term prognosis for a patient with fibromuscular dysplasia on an anti-platelet regimen? Until this month I wouldn't have understood the question or even been particularly interested. That was before the patient with this rare condition was my uncle and the person asking was my grandmother.

Unfortunately, the kind of research I do — fundamental aspects of yeast-cell division — doesn't provide the answers she's looking for. There are no side effects or case studies, and certainly no patients. Although I work on yeast cells because of their functional similarity to human cells, sheer curiosity is largely motivation enough. That has helped me understand a lot about yeast, but it doesn't do much for people like my uncle. I don't expect patients to run to their doctors saying, "I'm worried about the protein complexes forming at my origins of replication."

Having a person in my family deal with a disease that's scarcely mentioned in the medical literature has changed my perspective. It's not that I want to drop what I'm doing and take up applied research — I realize important contributions to humanity come from both directed and broad research approaches. I do, though, feel a new sense of meaning in what I do. The idea that my work may some day, however indirectly, reach a patient is no longer an abstraction but an aspiration my grandmother and I can share. ■

Milan de Vries is a molecular-biology graduate student at the Massachusetts Institute of Technology, Cambridge.

Take over

A real identity crisis.

Jon Courtenay Grimwood

It began when someone stole my wallet. Now, I'm not the kind of guy from whom you steal anything, and it was a good wallet — real alligator — so I was upset.

"Excuse me," he said. "I'm so sorry."

And standing up from where I'd seemingly tripped him in my hurry to get to the departures desk, he bowed. After another apology, he left in the opposite direction, still bowing.

He was foreign, obviously. Someone local would have called an attorney before even picking themselves off the floor.

At the departures desk, I put my hand inside my jacket and scowled. You know that look? The one you see in other people's eyes that says, why me? That was the look in the eyes of the boy on the other side of the desk.

"My wallet," I said.

He waited.

I searched my other pockets, in the way that you do. Even though you know it should have been in the first pocket, and if it isn't in that pocket, then it's not going to be in any of my other pockets either.

"It's gone..."

A security guard was moving towards me before the boy even had time to suggest that I leave the queue before taking another look.

"Problems, Sir?"

"My wallet," I told him. "Someone's stolen it."

He did that thing security guards do where he checks your clothes and watch and shoes, and for all I know your haircut and whether you have dirt under your nails. Whatever he saw, it was enough to put some politeness back into his voice.

"When did this happen?"

"Just now," I said. "When I was on my way to the desk. A pickpocket..."

We walked to the Ops Room together, he got me a paper cup full of water and then showed me to a chair. I'd done an iris scan at the door and a computer obviously matched this to the scan I did on arrival at the airport, because I came up on the screen immediately.

"That's you," he said. It was one of those half questions.

"Yeah," I said. "That's me."

I was the broad-shouldered man pushing his way through a crowd, impervious to the scowls around me. (I've never been good in crowds. I'm a guy who likes his space.)



JACEY

"Watch," I said.

The guard did as he was told. And he kept watching as I suddenly stumbled, looked round in a slightly drunken fashion and began talking to myself. The conversation lasted about 30 seconds, and finished with me giving an embarrassed little bow to no one in particular.

"That's impossible," I said, half rising from my seat.

"Let's take another look," said the guard. His hand on my shoulders could have been friendly. But when another two guards came into the room behind him I knew it wasn't. Actually, I'd known from the beginning.

All four of us watched the sequence.

Then we watched it again, from the moment I entered the airport to the moment it was my turn at the departures desk. A voice recorder had captured my conversation with the check-in clerk, and a box in the corner of the screen ran diagnostics. It registered 98% conviction in my voice.

"He's good," one of the guards said.

"I was robbed."

The guard shrugged towards the screen. "Doesn't look that way to me. You should have realized we'd have cameras everywhere." Glancing at the others, he raised his eyebrows.

"I've got to collect the kid," said one.

Another nodded.

"Toss him," said the guard who'd escorted me from the desk. "We can do without the grief."

Two of the guards walked to a back door, and one of them held me while the other punched me hard in the gut. He had muscles developed under full gravity. And the blow was hard enough to double me over. I'm not sure my feet even touched the three

steps it took to reach the ground.

"Don't come back," he said.

And that's my story... I did go back, of course. I went with a lawyer from one of those booths outside the departures lounge. He took my watch as surety and is probably still wearing it.

The guards denied ever having seen me before. As did the boy on the check-in desk, although less convincingly. They denied having punched me or shown me a replay of myself walking across the floor. And, unfortunately, the file for those few minutes had corrupted. So that was no use either.

The man living in my habitat has my ship, my bank account, my eyes, my DNA and, for all I know, my girlfriend. He seems to be doing a good job of running the corporation and I see shares are up again.

Obviously enough, he also has my wallet. From which he took all the information necessary to achieve this. So I'm waiting to go into court. Only accusing someone of identity theft is a serious matter. And the courts are not kind to vexatious accusations these days.

So, my question is, should I take the money? Should I drop my claim and sign an agreement never to repeat it? The sum I've been offered is large enough to see me out the rest of my life in relative comfort. And I won't be working 18-hour days or taking endless board meetings with people who want to replace me.

Because that's already been done. By someone so smart even my PA will probably never notice.

Jon Courtenay Grimwood is a full-time novelist. The first of his Ashraf Bey novels won the British Science Fiction Association award. His new novel is *End of the World Blues* (Gollancz).